

**MECHANISMS OF OBJECT REPRESENTATION IN INFEROTEMPORAL  
CORTEX**

by

Julianne E. Rollenhagen

Sc.B., Brown University, 1994

Submitted to the Graduate Faculty of  
Arts and Sciences in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

University of Pittsburgh

2003

UNIVERSITY OF PITTSBURGH  
FACULTY OF ARTS AND SCIENCES

This dissertation was presented

by

Julianne E. Rollenhagen

It was defended on

October 7, 2003

and approved by

Thomas D. Albright, Ph.D.

Marlene Behrmann, Ph.D.

Carol L. Colby, Ph.D.

Tai Sing Lee, Ph.D.

Carl R. Olson, Ph.D., Advisor

Anthony A. Grace, Ph.D.  
Committee Chairperson

# MECHANISMS OF OBJECT REPRESENTATION IN INFEROTEMPORAL CORTEX

Julianne E. Rollenhagen

University of Pittsburgh, 2003

The inferotemporal cortex in primates is thought to be the primary region that subserves object recognition. The studies presented here help to elucidate the role of IT in higher visual processing by addressing three specific outstanding issues. In the first study, we sought to determine whether IT neurons respond similarly to patterns that are perceptually confused. We considered a behavioral phenomenon whereby lateral mirror images are confused more frequently than vertical mirror images. By presenting mirror images to the monkey while simultaneously recording from IT neurons, we found that neurons differentiate less effectively between lateral mirror images than between vertical mirror images. This phenomenon may underlie the perceptual confusion documented in behavioral studies.

In the second study, we sought to determine whether activity in IT reflects experience-based changes in perception. We tested this by first training monkeys to discriminate shape orientation. We then recorded from IT neurons while monkeys performed an orientation discrimination task with trained orientations, and passively viewed orientations of trained and untrained shapes. We found that training to discriminate between orientations of a shape significantly increases the ability of IT neurons to discriminate between those same orientations. This neuronal selectivity

correlated with the monkeys' ability to discriminate orientation. These data suggest that training-induced changes in perception are supported by processes in IT.

Some IT neurons respond to the onset of a visual stimulus by firing a series of bursts at a frequency of around 5 Hz. One explanation for this phenomenon is that stimuli in the visual scene compete, with alternating success, for processing resources in IT. In the third study, we tested this by examining the oscillatory activity of IT neurons in response to the presentation of multiple stimuli, a central "preferred" image and a peripheral "non-preferred" image. We observed that the onset of a central pattern in the presence of the peripheral stimulus elicited strong oscillations phase-locked to pattern-onset. Onset of the peripheral stimulus in the presence of the central pattern elicited a succession of inhibitory troughs phase-locked to stimulus-onset. These results are congruent with a model of mutual inhibition of competing neuronal populations.

## FORWARD

### *Publications*

This dissertation is comprised of a general introductory chapter, followed by three chapters, each representing a separate research project, and a final general discussion chapter. The three research projects are listed below and have either been published or are being prepared for submission.

**Chapter 2:** Rollenhagen, J. E., & Olson, C. R. (2000). Mirror-image confusion in single neurons of the macaque inferotemporal cortex. *Science*, 287(5457), 1506-8.

**Chapter 3:** Rollenhagen JE and Olson CR. Effects of learning to discriminate shape orientation on neural responses in macaque inferotemporal cortex. *In preparation*.

**Chapter 4:** Rollenhagen JE and Olson CR. Low frequency oscillations arising from competitive interactions between visual stimuli in macaque inferotemporal cortex. *In preparation*.

## *Acknowledgements*

Carrying out this body of research was a huge endeavor and I am indebted to the many people who helped throughout the long process. First and foremost I must thank my advisor, Carl Olson. From spending countless hours teaching me the details of neurophysiological recording to helping me fine tune the finished product, he has provided continual guidance, without which the work in this thesis would never have been completed. I would also like to thank my committee members, Marlene Behrmann, Carol Colby, Tony Grace and Tai Sing Lee. They have served on many of my committees throughout my graduate career and have always provided invaluable feedback and guidance. In particular, I would like to extend a special thanks to Carol Colby for always being ready to help and advise me whenever I run across the hall and knock on her door. I would also like to thank both Carson Chow and Bill Eddy for their helpful advice on data analysis in Chapters 3 and 4.

I am grateful for the support of the many post docs, graduate students and technicians in the Colby and Olson labs. In particular, the completion of this work is in part due to the help of several technicians who have aided in training, data analysis and technical support over the years: Karen Medler, Jed Meltzer, Colleen Gault and Mike Handke. I especially owe a huge thanks to Karen McCracken, not only for providing technical support, but also for always being there to chat with and laugh with. Thanks also to Veerle Gysen for reading my thesis and offering very helpful comments.

It was only possible to complete the work described in this document because of the love and support of my friends and family. In particular, I would like to thank my friends, both in Pittsburgh and beyond for providing fun and necessary distractions to get

me out of the Mellon Institute! I would like to give a very special thanks to my parents, Dave and Debby Rollenhagen. They've inspired me, supported me and always had unwavering faith in me. Finally, I would like to thank Chris Baker, for being my colleague, mentor and, most of all, my best friend.

## TABLE OF CONTENTS

<b>Chapter 1: General Introduction .....</b>	<b>1</b>
1.1 General Overview .....	1
1.2 The Visual system and Two Main Processing Pathways .....	3
1.3 Inferotemporal Cortex and Object Processing.....	4
1.3.1 <i>Anatomical and Functional Subdivisions of IT</i> .....	5
1.3.2 <i>Role of IT in Visual Object Perception and Recognition</i> .....	7
1.4 The Neural Basis of Perceptual Confusion.....	10
1.5 Experience-Based Changes in IT.....	12
1.6 Oscillatory Visual Responses in IT .....	15
1.7 Goals .....	17
 <b>Chapter 2: Mirror Image Confusion in Macaque Inferotemporal Cortex .....</b>	<b>19</b>
2.1 Introduction.....	19
2.2 Methods .....	22
2.2.1 <i>Subjects</i> .....	22
2.2.2 <i>Stimuli</i> .....	22
2.2.3 <i>Behavioral Paradigm</i> .....	25
2.2.4 <i>Statistical Analysis</i> .....	26
2.3 Results.....	30
2.3.1 <i>Location of Recording Sites</i> .....	30
2.3.2 <i>Foveal Presentation</i> .....	30
2.3.3 <i>Time-course of Lateral Mirror Image Confusion</i> .....	35
2.3.4 <i>Peripheral Presentation</i> .....	36
2.3.5 <i>Mirror Image Preference in the Two Hemifields</i> .....	40
2.3.6 <i>Control for Symmetry</i> .....	41
2.4 Discussion.....	42
2.5 Summary .....	45
 <b>Chapter 3: Effects of Training to Discriminate Shape Orientation on Neural Responses in Macaque Inferotemporal Cortex .....</b>	<b>46</b>
3.1 Introduction.....	46
3.2 Methods .....	51
3.2.1 <i>Subjects</i> .....	51
3.2.2 <i>Stimuli</i> .....	51
3.2.3 <i>Training</i> .....	52
3.2.4 <i>Data Collection</i> .....	57
3.2.5 <i>Statistical Analysis</i> .....	58
3.3 Results.....	60
3.3.1 <i>Overview</i> .....	60



3.3.2 Location of Recording Sites .....	61
3.3.3 DMS Task: Behavior.....	62
3.3.4 DMS Task: Neural Activity .....	69
3.3.5 DMS Task: Correlation Between Behavioral Errors and Neuronal Activity ..	73
3.3.6 Fixation Task: Trained vs. Untrained Shapes .....	75
3.3.7 Fixation task: Trained vs. Untrained Orientations.....	85
3.3.8 DMS Task vs. Fixation task .....	93
3.4 Discussion.....	102
3.4.1 Overview .....	102
3.4.2 Effects of Training on Neuronal Response Strength.....	103
3.4.3 Effects of Training on Neuronal Selectivity for Shape Orientation .....	104
3.4.4 Extent of Selectivity Changes Across Neuronal Populations.....	106
3.4.5 Lateral Mirror Image Confusion in IT Neurons .....	106
3.4.6 Behavior in Relation to Neuronal Activity During Task Performance.....	107
3.4.7 Effects of Task Context on Neuronal Responses.....	108
3.4.8 Delay Period Activity.....	109
3.4.9 Behavioral Confusion of Lateral Mirror Images.....	111
3.5 Summary .....	112

#### **Chapter 4: Low frequency oscillations arising from competitive interactions between visual stimuli .....**

<b>between visual stimuli .....</b>	<b>114</b>
4.1 Introduction.....	114
4.2 Methods .....	116
4.2.1 Subjects .....	116
4.2.2 Stimuli .....	116
4.2.3 Lever Task.....	119
4.2.4 Fixation Task .....	120
4.2.5 Statistical Analysis .....	121
4.3 Results.....	128
4.3.1 Overview .....	128
4.3.2 Location of Recording Sites .....	130
4.3.3 Experiment 1: Example of Oscillatory Visual Response .....	130
4.3.4 Experiment 1: Model-Based Analysis.....	133
4.3.5 Experiment 1: Fourier Analysis.....	141
4.3.6 Experiment 2: Example of Oscillatory Visual Response .....	146
4.3.7 Experiment 2: Model-Based Analysis.....	148
4.3.8 Experiment 2: Fourier Analysis.....	152
4.3.9 Additional Observations .....	155
4.4 Discussion.....	160
4.4.1 Overview .....	160
4.4.2 Low-Frequency Oscillatory Activity in IT .....	160
4.4.3 Stimulus-Stimulus Interactions in IT.....	161
4.4.4 A Potential Mechanism .....	162
4.4.5 Relation to Biased Competition .....	163
4.5 Summary .....	165

<b>Chapter 5: General Discussion .....</b>	<b>167</b>
5.1 Summary .....	167
5.2 Neurons in IT Exhibit Lateral Mirror Image Confusion.....	168
5.3 Training Increases the Selectivity of IT Neurons .....	172
5.4 Low-Frequency Oscillations in IT .....	175
5.5 Conclusions.....	178
<b>Appendix: General Methods.....</b>	<b>180</b>
A.1 Introduction.....	180
A.2 Chair Training.....	180
A.3 Pedestal Implantation and Care .....	181
A.4 Placement of Recording Chamber .....	182
A.5 Electrophysiological Recording Methods.....	183
A.6 Data Display, Storage and Analysis.....	184
A.7 Localization of Recording Sites.....	184
<b>Bibliography .....</b>	<b>187</b>

## LIST OF TABLES

<b>Table 1.</b>	Significant differences in neuronal activity during the presentation of members of mirror-image pairs.	33
<b>Table 2.</b>	Lateral mirror image confusion analysis by ANOVA.	34
<b>Table 3.</b>	Behavior of monkeys during electrophysiological recording.	63
<b>Table 4.</b>	Numbers of neurons included in the model-based analysis.	132

## LIST OF FIGURES

<b>Figure 1.</b>	Dorsal and ventral streams.	3
<b>Figure 2.</b>	Coronal section through IT.	5
<b>Figure 3.</b>	Lateral and vertical mirror images.	19
<b>Figure 4.</b>	Stimuli used in the foveal and peripheral fixation tasks.	23
<b>Figure 5.</b>	Eight mirror image variants.	24
<b>Figure 6.</b>	Fixation task.	28
<b>Figure 7.</b>	Parasagittal magnetic resonance image of the right hemisphere of monkey Op.	28
<b>Figure 8.</b>	Example of a single neuron exhibiting lateral mirror image confusion.	29
<b>Figure 9.</b>	Lateral mirror image confusion for images presented at the fovea.	32
<b>Figure 10.</b>	Time-course of lateral mirror image confusion.	32
<b>Figure 11.</b>	Lateral mirror image confusion in neurons for images presented at the fovea and in the periphery.	39
<b>Figure 12.</b>	Stimuli used in training and data collection.	53
<b>Figure 13.</b>	Delayed match to sample (DMS) task used in training and data collection.	54
<b>Figure 14.</b>	Fixation task used in data collection.	55
<b>Figure 15.</b>	Examples of orientations of trained and untrained shapes presented in the fixation task.	56
<b>Figure 16.</b>	Coronal magnetic resonance images.	61

<b>Figure 17.</b>	Learning curves for each of ten shapes seen in the DMS task during training.	64
<b>Figure 18.</b>	Reaction times in the DMS task.	66
<b>Figure 19.</b>	Scan path of monkeys during search period of DMS task.	67
<b>Figure 20.</b>	Population histograms of neural activity collected during performance of the DMS task.	71
<b>Figure 21.</b>	Correlation between neuronal orientation selectivity and behavioral discrimination of orientation in the DMS task.	74
<b>Figure 22.</b>	Responses of a single neuron to orientations of trained and untrained shapes presented foveally in the fixation task.	76
<b>Figure 23.</b>	Comparison of neural activity elicited by orientations of trained and untrained shapes presented in the fixation task.	78
<b>Figure 24.</b>	Frequencies of orientation pairs for both trained and untrained tetrads that elicited significantly different responses from neurons.	83
<b>Figure 25.</b>	Comparison of neural activity elicited by trained and untrained orientations of trained shapes presented in the fixation task.	88
<b>Figure 26.</b>	Frequencies of orientation pairs for trained and untrained orientations that elicited significantly different responses from neurons.	92
<b>Figure 27.</b>	Comparison of neural activity elicited by trained orientations presented in the fixation task and DMS task.	96
<b>Figure 28.</b>	Correlation between neuronal responses for trained orientations presented in the DMS and fixation tasks.	100
<b>Figure 29.</b>	Frequencies of orientation pairs for trained orientations run in the DMS and fixation tasks that elicited significantly different responses from neurons.	101
<b>Figure 30.</b>	Data collection tasks.	117
<b>Figure 31.</b>	Lever response in Experiment 1.	118
<b>Figure 32.</b>	Example of curve-fitting from model-based analysis.	123

<b>Figure 33.</b>	Example neurons from Experiments 1 and 2.	131
<b>Figure 34.</b>	Measures of oscillatory activity.	134
<b>Figure 35.</b>	Population histograms.	137
<b>Figure 36.</b>	Frequency of oscillations.	139
<b>Figure 37.</b>	Mean amplitude of oscillatory activity.	140
<b>Figure 38.</b>	Time-constants of oscillatory decay.	142
<b>Figure 39.</b>	Auto-correlograms (ACGs) and power spectra of data from the two example neurons shown in Figure 33.	144
<b>Figure 40.</b>	Population power spectra.	145
<b>Figure 41.</b>	Within neuron comparisons of spectral power in the frequency range of 4-7 Hz.	147
<b>Figure 42.</b>	Population histogram representing mean firing rate as a function of time under the flanker-alone condition and the object-then-flanker condition for all neurons exhibiting oscillatory visual responses under the latter condition.	156
<b>Figure 43.</b>	Example of a neuron more responsive to flanker.	157
<b>Figure 44.</b>	Data from a neuron giving an initially positive oscillatory response to the object presented in isolation (top row) and an initially negative oscillatory response to the flanker presented against the backdrop of the already present object (bottom row).	159
<b>Figure 45.</b>	A simple model of oscillatory activity arising from reciprocal inhibition between neurons selectively responsive to different visual stimuli and subject to fatigue.	164

# **Chapter 1**

## **General Introduction**

### **1.1 General Overview**

As we move through the world, we are constantly taking in visual information about our surroundings. The visual patterns that fall onto our retinas are continuously being processed by our nervous system, allowing us to recognize a friend on the street, discriminate between the flowers in our garden and find our car in a crowded parking lot. Such seemingly simple tasks are the result of highly complex neural processing involving numerous interconnected regions of the brain. Yet, despite recent advances in our knowledge of cortical visual processing, the mechanisms involved are poorly understood, and current artificial visual systems lag far behind human performance. Elucidating how these brain regions process visual patterns will not only aid in understanding the neural basis of complex behavior but will also have important implications in developing object recognition algorithms for artificial vision and revealing consequences of visual cortical dysfunction. The goal of the studies presented in this thesis is to gain a better understanding of the neural mechanisms underlying pattern processing. Specifically, the experiments reported here aim to shed light on the role of the inferotemporal cortex (IT) in macaque monkeys, in the processing of complex shapes. Three specific issues related to processing in IT that remain unresolved will be addressed in these studies. The first issue concerns whether the activity in IT reflects the degree to which complex shapes are perceived as similar. In particular, we demonstrate in Chapter 2 that IT neurons

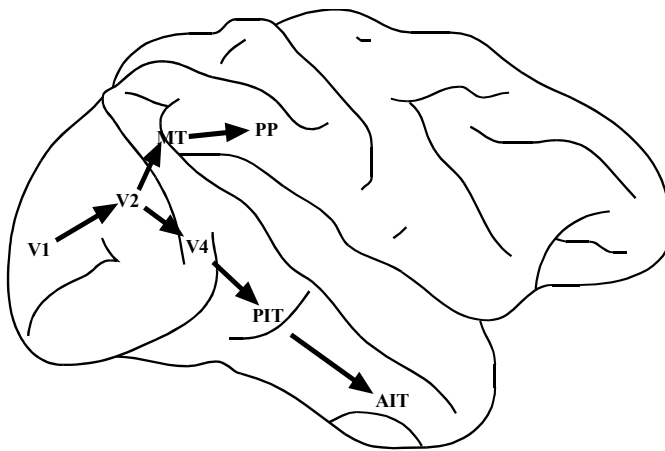
differentiate less effectively between lateral mirror images than vertical mirror images, a phenomenon that parallels the perceptual confusion shown in behavioral studies. The second issue concerns whether experience-based changes in perception are reflected in the responses of IT neurons. In Chapter 3, we demonstrate that orientation discrimination training leads to significant increases in the selectivity of IT neurons for trained orientations. Finally, the third issue addressed in this thesis concerns the nature of low-frequency oscillations that are often observed in the responses of IT neurons. In Chapter 4, we characterize these oscillations and demonstrate that they are enhanced when a second stimulus is added to the visual scene.

The goal of the present chapter is to give a brief overview of IT to illustrate the generally accepted view that this region is involved in object recognition. Further, this chapter will provide general background for the experiments described in Chapters 2-4. More detailed background for these studies will be provided in the individual chapters. Specifically, I will begin with a brief introduction to the organization of the primate visual system, and then focus on IT cortex, providing an overview of its role in pattern processing by discussing its subdivisions, cell properties, and connectivity. I will then discuss, in turn, the role of IT in encoding perceptual similarity, learning-induced changes in IT, and evidence for oscillatory activity in IT, thus providing general background for the experiments outlined in Chapters 2-4. Finally, I will conclude by laying out the specific goals of these three studies.



## 1.2 The Visual system and Two Main Processing Pathways

On the basis of both lesion and anatomical studies, cortical visual processing in both human and non-human primates is thought to be organized into two major pathways (Figure 1). The dorsal visual stream, often termed the “where” (Ungerleider and Mishkin, 1982), or “how” (Milner et al., 1991; Goodale and Milner, 1992; Milner and Goodale, 1995) pathway, extends from primary visual, or striate cortex (V1) through V2, V3 and MT to regions within the posterior parietal cortex (Ungerleider and Haxby,



**Figure 1. Dorsal and ventral streams.** Dorsal stream extends from primary visual cortex (V1) through V2, V3 and MT to regions within the posterior parietal cortex. Ventral stream extends from primary visual cortex through V2 and V4 to regions in the inferotemporal lobe.

1994). Results of lesion and neurophysiological studies suggest that this pathway is concerned primarily with the processing of information necessary for carrying out such actions as reaching for an object or making eye movements to a location in space. For example,

lesions to this pathway result in severe deficits in spatial perception (Ungerleider and Mishkin, 1982). Moreover, neurons in dorsal stream regions have been shown to respond to visual stimuli (Colby and Duhamel, 1996; Murata et al., 2000) and in concert with eye movements (Colby and Duhamel, 1996) and arm movements (Snyder et al., 2000) to these stimuli.

The ventral visual stream, or the “what” pathway (Ungerleider and Mishkin, 1982; Goodale and Milner, 1992), extends from V1 through V2 and V4 to regions in the

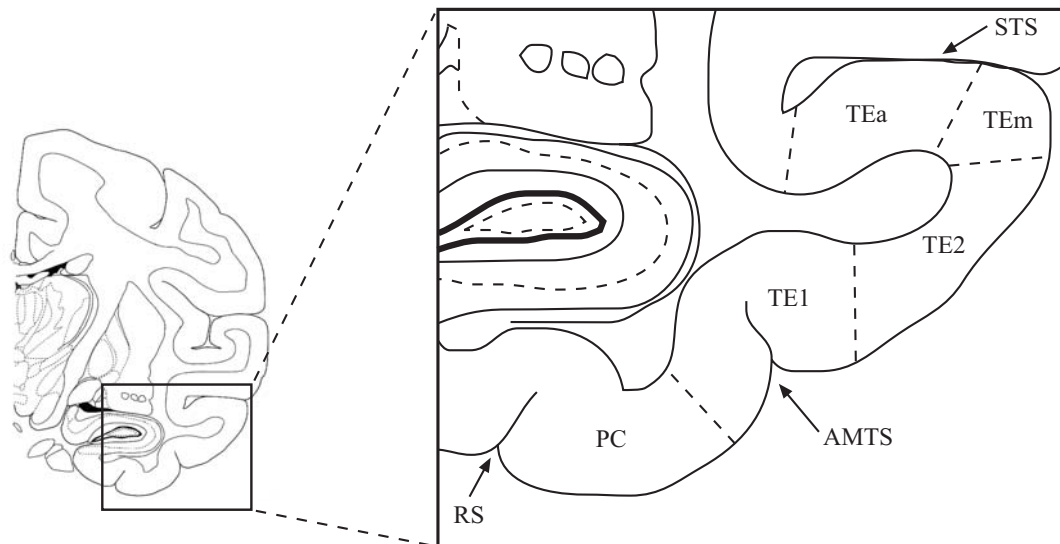
inferior temporal lobe (Ungerleider and Haxby, 1994). This pathway is thought to be concerned primarily with the processing of complex patterns or objects, for the purposes of object perception and recognition. This is supported by studies demonstrating that lesions of this pathway result in impaired performance in the discrimination and recognition of complex objects (Covey and Gross, 1970; Dean, 1976; Ungerleider and Mishkin, 1982). Additionally, neurons in ventral stream regions respond selectively to complex visual stimuli (Tanaka et al., 1991; Pasupathy and Connor, 2001). At the end of this ventral visual stream lies the inferotemporal cortex, or, IT, the final purely visual area that is thought to be strongly involved in object perception and recognition.

### **1.3 Inferotemporal Cortex and Object Processing**

IT is a general term referring to the large region of cortex that lies along the ventral aspect of the temporal lobe. Specifically, IT is considered to be that cortex which lies anterior to V4 and ventral to MT, FST, MST and STP, encompassing cortex from the fundus of the superior temporal sulcus (STS) ventrally to the rhinal and occipitotemporal sulci, and extending rostrally to the end of the superior temporal sulcus (Figure 2). IT itself is thought to encompass several sub-regions, although the nomenclature and definition of these regions varies extensively within the literature. I will begin by reviewing some of the terminology and principles by which the subdivisions have been determined before describing the functional properties of this region.

### 1.3.1 Anatomical and Functional Subdivisions of IT

In general, areas within IT have been defined both anatomically and functionally. Evidence for anatomical subdivisions, however, is weak and contradictory. Based on connectivity, Felleman and Van Essen (1991) divided IT into six main regions, the dorsal and ventral areas of PIT, CIT and AIT (posterior, central and anterior inferotemporal cortex). In contrast, IT has been divided in terms of cytoarchitectonics into two main regions, TEO, located posteriorly, and TE, located anteriorly (von Bonin and Bailey, 1947). In addition, there is a region that lies medial to the anterior medial temporal sulcus (AMTS) that is cytoarchitectonically distinct from TE, termed perirhinal cortex (corresponding to Brodmann's areas 35 and 36; Saleem and Tanaka, 1996). Finally, subdivisions of TE defined on the basis of cytoarchitecture and connectivity include, in order, from the lower bank of the STS to the AMTS, TEa, TEm and TE<sub>1-3</sub> (Seltzer and Pandya, 1978).



**Figure 2. Coronal section through IT.** IT extends from the fundus of the STS ventrally to the rhinal and occipitotemporal sulcus (not shown). Anatomical subdivisions defined by Seltzer and Pandya (1978) are shown. RS: rhinal sulcus; PC: perirhinal cortex; AMTS: anterior medial temporal sulcus; STS: superior temporal sulcus.

Evidence for functional subdivisions within IT is also weak. For example, Baylis et al. (1987) recorded from each of the subdivisions defined cytoarchitectonically by Seltzer and Pandya (1978), but found little difference in response properties between the areas. While Janssen et al. (2000) found subtle differences in selectivity for three-dimensional shapes between the lower bank of STS (which corresponds to cytoarchitectonic areas TE<sub>a</sub> and TE<sub>m</sub>) and the lateral convexity, the two areas were found to be equal in selectivity for two-dimensional shape. Tamura and Tanaka (2001) observe that there was greater activation of neurons by colorful stimuli in TE<sub>av</sub> (which corresponds roughly to area TE<sub>1</sub>) than in TE<sub>ad</sub> (which corresponds roughly to area TE<sub>2</sub>), but no statistical comparison of the two areas was carried out. In contrast to the smaller cytoarchitectonic distinctions, the larger anatomical subdivisions are marked by subtle but definite differences in functionality. In particular, IT has been functionally subdivided into two primary areas, anterior IT (AIT) and posterior IT (PIT), which roughly correspond to the anatomically defined TE and TEO, respectively. This division is based on findings that both the receptive field size and the complexity of stimuli necessary to drive neurons gradually increases as one moves further anterior (Tanaka et al. 1991; Kobatake and Tanaka, 1994). In addition, there is some evidence that mnemonic encoding is greater in perirhinal regions than in TE (Naya et al., 2003), although, for the most part, neurons in both areas exhibit similar selectivity for complex stimuli.

The strongest evidence for regional specialization in IT concerns organization on a more local (columnar) scale. Specifically, neurons with similar feature selectivity have been shown to cluster together (Tanaka et al. 1991). By recording in vertical penetrations

through IT cortex, Fujita et al. (1992) found that these clusters form columns perpendicular to the surface of the cortex. These columns contain neurons that respond to similar features and have thus been termed “feature columns”. Further, Tsunoda et al. (2001) used a combination of optical imaging and electrophysiological recordings to demonstrate that an object was represented in a spatially distributed manner, activating several feature columns, each responding to a specific visual feature of the object.

The studies reviewed above offer little evidence for either functional or anatomical subdivisions within IT. Therefore, I will now refer to IT in general, as essentially the region referred to as TE.

### *1.3.2 Role of IT in Visual Object Perception and Recognition*

There are several lines of evidence supporting the idea that the inferotemporal cortex is involved in the perception and recognition of complex objects. One line of evidence comes from lesion studies in monkeys. Several groups have demonstrated that monkeys with inferotemporal lesions have basic sensory capabilities intact, but are profoundly impaired in tasks requiring visual pattern recognition (Cowey and Gross, 1970; Dean, 1982).

A second line of evidence comes from single-neuron recording experiments in monkeys. These experiments have demonstrated that IT neurons possess properties consistent with their mediating object perception and recognition. The most prominent of these properties is selectivity for particular complex shapes. Numerous studies have demonstrated that neurons in IT respond selectively to complex two-dimensional stimuli, including faces (Gross et al., 1972; Perrett et al., 1982; Desimone et al., 1984; Baylis and

Rolls, 1987; Yamane et al., 1988; Tanaka et al., 1991), hands (Gross et al., 1969, 1972; Desimone et al., 1984; Tanaka et al., 1991) and inanimate objects such as brush-like stimuli and fourier descriptor stimuli (Gross et al., 1969 and 1972; Schwartz et al., 1983; Desimone et al., 1984; Albright and Gross, 1990; Tanaka et al., 1991). IT neurons have also been shown to be selective for disparity-defined three-dimensional shapes (Janssen et al., 1999). Furthermore, several groups have demonstrated that the selectivity of IT neurons for complex stimuli is maintained across changes in location (Schwartz et al., 1983; Sary et al., 1993; Ito et al., 1995), size (Sato et al., 1980; Schwartz et al., 1983; Sary et al., 1993; Ito et al., 1995), defining cue, such as texture or motion (Sary et al., 1993), and partial occlusion (Kovacs et al., 1995). This is consistent with the fact that at a behavioral level, recognition is invariant across these changes.

Another distinguishing characteristic of IT neurons is their possessing large receptive fields which almost always include the fovea and usually extend into the opposite hemifield. Most neurons have receptive fields that are over 10x10 degrees of visual angle, many over 30x30 degrees (Gross et al., 1969). It is thought that neurons in earlier regions of cortex with more restricted receptive fields converge onto neurons in IT, allowing for the activation of neurons by preferred objects regardless of precise location. Exceptions to this rule do exist. Recent studies have identified receptive fields in anterior IT as small as 3° in diameter (Op de Beeck and Vogels, 2000). Also, the precise size of IT receptive fields seems to depend on the stimulus size, with larger stimuli eliciting visual responses over a broader range of positions (Op de Beeck and Vogels, 2000).

Finally, a third line of evidence supporting the involvement of IT in the perception and recognition of complex objects comes from examination of the connectivity of IT with other brain regions. Specifically, IT receives input from earlier cortical regions within the ventral stream that relay visual information from the retina, and projects to areas of higher order involved in memory encoding and high level cognitive functioning, including executive control. For the most part, visual information is passed to IT through a serial pathway extending from primary visual cortex through V2, V4, posterior IT and, finally, anterior IT (Desimone et al., 1980). Exceptions to this strict serial connectivity include feed-forward projections that connect V2 to TEO (Nakamura et al., 1993), and V4 to the posterior portion of TE (Desimone et al., 1980). IT is strongly linked to several regions of cortex implicated in supravisual cognitive processes. These include prefrontal cortex (primarily areas 12 and 45, with minor projections to areas 8, 11 and 13; Webster et al., 1994), the amygdala (Herzog and Van Hoesen, 1976; Aggleton et al., 1979; Cheng et al., 1997), the neostriatum (Webster et al., 1993; Cheng et al., 1997), parietal area LIP (Webster et al., 1994), and parahippocampal area TF (Suzuki, 1996). Finally, sub-regions within IT itself are highly interconnected. Both TE and area 36 in perirhinal cortex are reciprocally connected with STS (Saleem et al., 2000), and there are extensive connections between TE and perirhinal cortex (Van Hoesen and Pandya, 1975). There also exist numerous connections within TE, between TE and TEO and between the inferotemporal cortices in the two hemispheres (Desimone et al., 1980).

The evidence from lesion, electrophysiological and anatomical studies described above supports the involvement of IT in late stages of visual processing underlying

perception and recognition. Yet, despite this evidence, many issues concerning the mechanisms of object processing remain unresolved. In the following sections, I will discuss, in turn, three outstanding questions in the field of inferotemporal research. First, I will explore the question of whether responses in IT neurons reflect the degree to which complex stimuli are perceived as similar, or simply their physical similarity. In particular, the extent to which mirror images are perceived as similar will be discussed and experiments examining responses to these images will be proposed. Next, I will examine evidence for changes in neuronal responses in IT resulting from object recognition and discrimination learning. Experiments will be proposed to reconcile conflicting results obtained to date. Finally, the time-course of neuronal signals in IT will be discussed. In particular, I will put forth evidence for low-frequency oscillatory activity in IT neurons and propose experiments to explore and quantify this phenomenon.

#### **1.4 The Neural Basis of Perceptual Confusion**

As described above, neurons in IT are selective for complex objects. It is unclear, though, whether these neurons encode stimuli simply in terms of their physical attributes, or whether their firing reflects the manner in which stimuli are perceived. Physical similarity is not always perfectly correlated with perceived similarity. For example, there is substantial evidence that lateral mirror images (images reflected over the vertical axis) are perceived as more similar than vertical mirror images (images reflected over the horizontal axis), despite the fact that lateral and vertical mirror images are equally similar with respect to physical attributes (Sutherland, 1960; Todrin and Blough, 1983; Hamilton and Tieman, 1973; Riopelle et al., 1964; Bornstein et al., 1978; Huttenlocher, 1967;



Serpell, 1971; Rudel and Teuber, 1963; Sekuler and Houlihan, 1968). Certainly for the most part, though, stimuli that are physically similar (as determined by some objective computational measure) are perceived as similar (as determined by some behavioral measure). For example, Sugihara et al. (1998) examined the perception of shape similarity for computer-generated animal-like stimuli that varied systematically in a high dimensional parameter space. Using error patterns from a delayed match-to-sample task, they found that the perception of similarity in non-human primates mirrors the pattern of similarity between objects as measured by parameter space distances. A similar experiment in humans demonstrated that humans also perceptually represent complex objects in terms of their parametric similarities (Cutzu and Edelman, 1998). Furthermore, non-human and human primates tend to represent the similarity between stimuli in a congruent manner (Tomanaga and Matsuzawa, 1992; Op de Beeck et al., 2001).

The role of IT in encoding perceived similarities has been addressed in two recent studies. In the first, Op de Beeck et al. (2001) compared the ability of monkeys to discriminate parameterized shapes and the selectivity of IT neurons for those same shapes. Monkeys were trained to determine whether pairs of stimuli were the same or different. The degree to which monkeys confused a pair of stimuli was taken as a measure of the similarity of those stimuli. The perceptual confusion of stimuli was highly congruent with the similarity of stimuli as coded by IT neurons. In this study, stimuli were used that were parametrically altered such that the perceived similarity was highly correlated with the physical similarity. Although the behavioral and neural representations deviated consistently from the parametric configurations, this effect was

subtle and leaves open the question of whether the activity of neurons reflects the physical attributes of the stimuli and not how they are perceived. Stimuli that were not specifically designed to be physically alike were incorporated into a comparable study designed to examine the selectivity of IT neurons for images that were perceived as similar (Miyashita et al., 1993). The stimulus set consisted of 97 fractal images that were rated for similarity by human observers. The results showed that there was a tendency for neurons to respond comparably to stimuli that were rated as highly similar by humans. Although the shape complexity of the stimuli was not parametrically varied, one cannot rule out the possibility that the stimuli in this study that were perceived as more alike were in fact physically similar. Thus, these studies do not provide conclusive evidence that neurons in IT encode perceived similarity.

One way to circumvent the issue of physical similarity is to consider the specific case of mirror image confusion described above. In this case, the perception of similarity differs for stimuli that are equal in their physical similarity by any isotropic measure. Therefore, this phenomenon allows one to ascertain the degree to which IT neurons confuse stimuli that primates confuse, without the confounding issue of physical similarity. In Chapter 2, experiments directly testing whether IT neurons confuse lateral mirror images more than vertical mirror images will be described.

### **1.5 Experience-Based Changes in IT**

Since IT is considered to be the primary site for object recognition and perception, it seems reasonable that experience-based changes in recognition and perception would rely on this region of cortex as well. In particular, as experience with objects increases,

accuracy in discriminating and recognizing these objects increases. Regions of cortex concerned with the coding of objects must be plastic to some degree to encode these changes. While much has been done to better our understanding of how visual information is represented in the activity of IT neurons, relatively little is known about the neural mechanisms underlying experience-based changes in perception.

IT is in a unique position to mediate these changes in perception. It receives visual information from earlier visual areas, and projects to areas of higher order critical for mnemonic function such as medial temporal, limbic and frontal cortex. Moreover, ablation studies have demonstrated that animals with IT lesions learn visual discriminations more slowly than normal animals, suggesting that IT is essential for normal visual discrimination learning (Dean, 1976).

If visual learning-induced changes occur in IT, it might be expected that these changes would be manifested in changes in the activity patterns of single IT neurons. Indeed, evidence for such changes has been demonstrated in IT under a number of different circumstances. For example, it has been shown that (1) repeated exposure to an initially novel stimulus leads to a decline in response strength (Rolls et al., 1989; Miller et al., 1991, Xiang and Brown, 1998) and changes in stimulus selectivity (Rolls et al., 1989), (2) training on a visual paired-associate task results in the emergence of neurons that are responsive to both members of a pair (Sakai and Miyashita, 1991; Erickson and Desimone, 1999; Messenger et al., 2002), (3) training to discriminate complex stimuli leads to increased selectivity (Logothetis and Pauls, 1995; Kobatake et al., 1998; Baker et al., 2002) and response strength (Kobatake et al., 1998) for trained compared to untrained stimuli, and increased selectivity for diagnostic features compared to non-diagnostic

features (Sigala and Logothetis, 2002), and (4) prior experience with stimuli leads to spatial clustering of neurons selective for those stimuli (Erickson et al., 2000). The time-course of these learning effects is not clear. While most studies have examined effects of long term training (Logothetis and Pauls, 1995; Kobatake et al., 1998; Baker et al., 2001), there is evidence to suggest that learning may occur on a shorter time frame (Erickson et al., 2000; Messenger et al., 2002).

Despite the numerous studies cited above showing learning effects in IT, experiments examining the impact of discrimination training on activity in IT have found weak and contradictory effects. In particular, there is little consensus on whether training affects the response strength, the selectivity of neurons, the percentage of neurons involved in encoding stimuli or a combination of all three. This lack of consensus may result from differences in experimental design and task requirements. For example, Kobatake et al. (1998) demonstrated that shape discrimination training leads to increases in response strength for trained stimuli. However, in this study, comparisons were made across trained and untrained animals, and therefore any effects observed could be due to inter-animal differences. In terms of selectivity, Logothetis et al. (1995) claimed that training increased the selectivity of a small subset of IT neurons for trained stimuli, but failed to provide quantitative analyses to support this claim. In contrast, Baker et al. (2001) demonstrated that IT neurons in monkeys trained on a feature conjunction task did not differ in firing rate for trained and untrained stimuli. Furthermore, these authors confirmed that training increased the selectivity of neurons for trained stimuli, but showed that this occurred as a subtle shift in selectivity across many neurons and not a strong shift in few.

While Baker et al. carried out a comprehensive, quantitative analysis of the effects of training on neuronal activity in IT, it is unclear to what degree these results are dependent on the specifics of the task design. Monkeys in this experiment were trained on a feature conjunction task which required attending to top and bottom features of unique baton-like stimuli. Therefore, it remains a question whether these effects would be observed under different task conditions. In Chapter 3 we address this question by examining the effects of orientation discrimination training on the activity of neurons in IT.

## **1.6 Oscillatory Visual Responses in IT**

In recent years there has been an interest in the time-course of neural signals in IT. Specifically, research has centered on investigating (1) the amount of information in the temporal pattern of responses and (2) the nature of oscillations in visual activity. The amount of information encoded in the spike trains elicited by IT neurons is not clear. For the most part, information carried by IT neurons about images seems to be in the form of a rate code: the spike count throughout the response period is greater in the presence of a preferred image than of a non-preferred image. However, the temporal pattern of the response can also vary across stimuli, suggesting that a rate code alone is not the only source of information (Richmond and Optican, 1987; Richmond et al. 1987). The importance of temporal codes was questioned by Miller et al. (1993a) who found that most of the information about a stimulus may be accessed via the average firing rate of the neurons, although there is some improvement in coding when considering the time-course of the signal. Likewise, Tovee et al. (1993) showed that a substantial portion of

the information transmitted by IT cells can be extracted from intervals as short as 50ms, and that the most information is in the initial portion of the visual response. In contrast, it has been suggested that different types of information are encoded at different times in the visual response. In examining responses of IT neurons to faces and shapes, Sugase et al. (1999) found that global information, categorizing stimuli as monkey faces, human faces or shapes, is encoded in the earliest part of the visual response. Finer information about the face identity or expression is encoded later in the response, beginning approximately 50ms after global information.

In contrast to temporal codes, few have investigated oscillatory activity in IT. While the responses of IT neurons are typically characterized by an initial transient burst (for 100-200 ms) followed by a lower, more maintained discharge (Oram and Perrett, 1992; Tamura and Tanaka, 2001), some responses consist of repeated bursts of firing at frequencies around 5 Hz. Although these low frequency oscillations are often observed in studies of IT responses, only one group to date has examined them in any depth. Nakamura and colleagues (1991, 1992) described oscillatory activity evoked in neurons in the temporal pole by complex visual stimuli such as pictures of faces, food and inanimate objects, during a visual discrimination task. While the frequencies of these oscillations varied slightly, the most common were between 5-6 Hz. It is unclear what the functional significance of this low-frequency oscillatory activity is. Nakamura et al. found that this pattern of activity was more commonly elicited by familiar objects and suggested that this is due to a heightened attention to, or an increased behavioral relevance of, the stimuli. However, oscillatory activity has never been cited as a

characteristic of learning-induced changes in IT (Rolls et al., 1989; Logothetis and Pauls, 1995; Kobatake et al., 1998; Baker et al., 2002).

Models of neuronal networks have demonstrated that oscillations may arise in neurons subject to fatigue and inhibitory input from other neurons (Wilson et al., 2000). Therefore, it is possible that these oscillations arise from competitive interactions between separate populations of neurons. It might be the case, then, that stimuli evoking responses in these antagonistic populations would lead to the oscillatory activity observed. Experiments described in Chapter 4 are designed to more quantitatively characterize low-frequency oscillations in IT and to determine under what stimulus conditions they occur.

## **1.7 Goals**

In summary, data from lesion and neurophysiological studies support the idea that IT is important in the processing of complex objects for perception and recognition. The studies that are laid out in this document aim to further examine the role of IT in object processing. We employed the method of recording from single neurons in IT of the awake behaving monkey to address the following issues.

In Chapter 2, we describe research aimed at determining whether the selective activity of IT neurons reflects the degree to which visual stimuli are perceived as similar. In particular, we investigated whether lateral mirror images evoke more similar responses from IT neurons than vertical mirror images. We studied the responses of IT neurons during the passive presentation of mirror image stimuli both at the fovea and in the

periphery to test the hypothesis that IT neurons differentiate less effectively between a pattern and its lateral mirror image than between a pattern and its vertical mirror image.

In Chapter 3, we describe research aimed at clarifying the nature of changes in visual responsiveness induced in IT by visual discrimination training. We initially trained monkeys to discriminate among four orientations of each of ten images. We then recorded from IT neurons while monkeys passively viewed trained and untrained images. Using these methods, we evaluated whether orientation discrimination training resulted in changes in the response properties of IT neurons.

In Chapter 4, we describe research aimed at characterizing the low-frequency oscillations observed in IT neurons and determining the conditions in which this pattern of activity arises. We examined the activity of IT neurons in response to the presentation of multiple stimuli, a central image that excited the neuron and a peripheral image that did not. We used these methods to test the hypothesis that these oscillations arise from competing populations of neurons responsive to different stimuli in the visual scene.

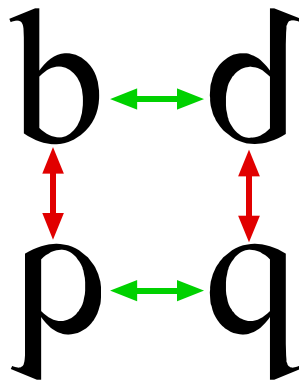


## Chapter 2

### Mirror Image Confusion in Macaque Inferotemporal Cortex

#### 2.1 Introduction

Some pairs of visual objects have greater perceptual similarity than others. This generalization applies in particular to images related by mirror-image transformation. Specifically, it has been demonstrated that lateral mirror images (Figure 3, green arrows) are perceived as more similar (as measured by the degree to which they are confused) than vertical mirror images (Figure 3, red arrows). The relative confusion of lateral



**Figure 3. Lateral and vertical mirror images.** Green arrows indicate lateral mirror image pairs and red arrows indicate vertical mirror image pairs.

mirror images over vertical mirror images has been termed lateral mirror image confusion. This phenomenon has been demonstrated in many species, including octopi (Sutherland, 1960), pigeons (Todorin and Blough, 1983), monkeys (Hamilton and Tieman, 1973; Riopelle et al., 1964) and humans (Bornstein et al., 1978; Huttenlocher, 1967; Serpell, 1971; Rudel and Teuber, 1963; Sekuler and Houlihan, 1968). For example, Riopelle et al. (1964) demonstrated that

monkeys trained to discriminate between two simultaneously presented line figures made more errors when those figures were lateral mirror images of each other than when they were vertical mirror images or different images. In humans, mirror image confusion has been demonstrated in adults (Sekuler and Houlihan, 1968), children (Huttenlocher, 1967;

Serpell, 1971; Rudel and Teuber, 1963) and infants (Bornstein et al., 1978). A study by Bornstein et al. (1978) used a preferential looking paradigm to show that human infants as young as 4 months old also perceive lateral mirror images as more similar to each other than vertical mirror images.

Several theories have been posed to explain why lateral mirror images are confused more than vertical mirror images. One theory suggests that the confusion of lateral mirror images arises as an accidental consequence of the bilateral symmetry of the nervous system. To the degree that the hemispheres are mirror images of each other and inter-hemispheric pathways are precisely symmetric, neurons in the left hemisphere activated by a 'b', for example, must be linked to neurons in the right hemisphere activated by a 'd', with the consequence that either stimulus will activate both populations, giving rise to confusion (Corballis and Beale, 1970).

A second theory is based on the fact that lateral reversals usually result from changes in viewpoint and therefore convey little information that is important for object recognition – a cup is still a cup whether the handle is on the right or left. Vertical reversals, however, rarely come about from changes in viewpoint and therefore do convey important information about the object. If lateral reversals convey little information, then the brain resources dedicated to representing them may have become relatively limited, through an adaptive phylogenetic or ontogenetic process (Gross and Bornstein, 1978).

Whichever account is true, the question still remains: does lateral mirror image confusion have a demonstrable neural correlate? Given that IT cortex is critical for object recognition and neurons in this region respond selectively to complex visual

stimuli, it is plausible that processes within this region may provide the neural underpinnings for lateral mirror image confusion. In particular, it might be the case that overlapping populations of neurons are activated by lateral mirror images, leading to the perceived similarity and the resulting behavioral confusion of these images. The idea that perceptual confusion, or perceived similarity, might arise from the responsiveness of neurons in IT has found support in a recent study by Op de Beeck et al. (2001) which demonstrated that neurons in IT respond more similarly to stimuli that monkeys find perceptually confusing.

If, behaviorally, lateral mirror images are perceived as more similar to each other than vertical mirror images, then it might be expected that neurons in IT would respond more similarly to lateral mirror images than to vertical mirror images. While this idea has not been directly tested, there are examples in the literature supporting the idea that lateral mirror images evoke similar responses in IT neurons. For example, some neurons in the banks of STS responding selectively to views of faces rotated in depth have bimodal tuning curves such that they respond best to the right and left profiles (Perrett et al. 1991). This pattern has also been observed in neurons responsive to non-face stimuli as well. For example, Logothetis et al. (1995) demonstrated that some neurons responsive to a particular view of a wire object also respond to the pseudo-mirror image of that view formed by rotating the object in depth. A further example comes from Tanaka et al. (1991), who showed a neuron that responded selectively to a bilaterally symmetrical image rotated in the plane (bilaterally symmetrical stimuli rotated 180° from each other form mirror images). This neuron responded equally well to lateral mirror images, but less so to vertical mirror images (their Figure 7E-H).

While these few examples support the idea that neurons in IT may provide the neural basis for the perceptual confusion of lateral mirror images, no study to date has carried out a systematic comparison of the selectivity of IT neurons for lateral and vertical mirror images. In the present study, we attempted to resolve this by recording from IT neurons during the presentation of both lateral and vertical mirror image stimuli. We found that IT neurons responded more similarly to lateral mirror images than to vertical mirror images, supporting the idea that a neural correlate of lateral mirror image confusion may reside in IT.

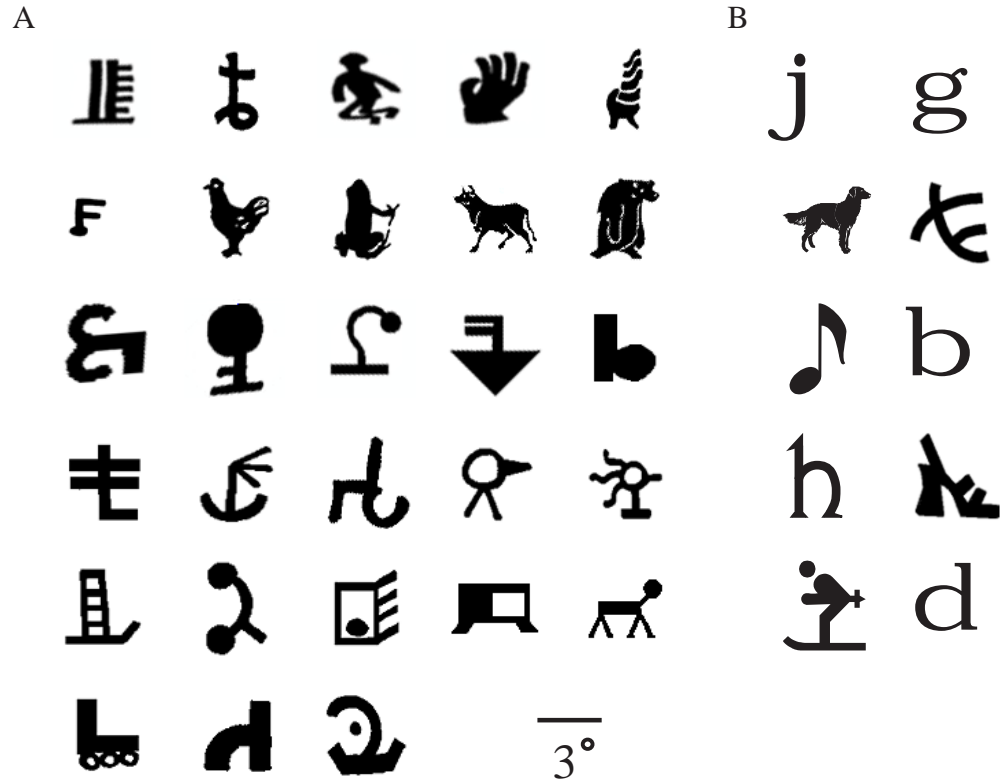
## **2.2 Methods**

### *2.2.1 Subjects*

Three adult male rhesus monkeys (*Macaca mulatta*) weighing 5.5-7.5 kg were used in this experiment. Their laboratory designations were Op, Fi and Ph. General surgical and training procedures are described in Appendix A. Training specific to this experiment is described below.

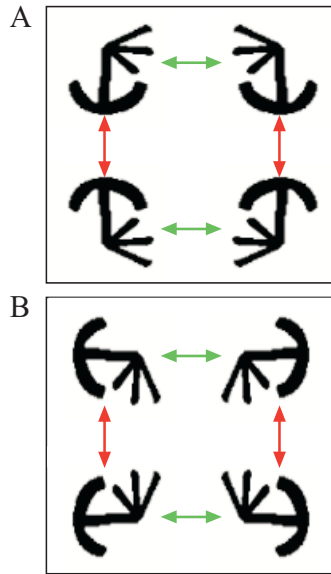
### *2.2.2 Stimuli*

The stimuli used in this experiment were thirty-eight white, asymmetric shapes each subtending approximately 3° of visual angle in height and width (Figure 4). Twenty-eight of these shapes were used with monkeys Op and Fi (Figure 4A). Ten were used with monkey Ph (Figure 4B). One of these shapes was selected for use in the main data collection tasks in the following manner. At the beginning of each recording



**Figure 4. Stimuli used in the foveal and peripheral fixation tasks.** A. Twenty-eight chiral shapes ( $3^\circ$  in height and width) used with monkeys Op and Fi. Shapes were presented one at a time, in white, against a black background, to the fixating monkey. One shape that elicited a strong response from the isolated neuron was identified for further use in data collection. B. Ten shapes used with monkey Ph.

session, a neuron in IT was isolated and each shape in the set was presented one at a time, foveally, to the fixating monkey. One shape was identified that elicited a strong response from the isolated neuron, as assessed by evaluating the audio monitor output and online histograms. If none of the shapes elicited a clear response from the neuron, the neuron



**Figure 5. Eight mirror image variants.** A. Four mirror image variants were formed by rotating the base shape 180° laterally in depth, 180° vertically in depth and 180° within the plane. B. A second tetrad of mirror images was created by performing the same transformations on the 90° rotation of the base shape. Green arrows indicate lateral mirror image pairs. Red arrows indicate vertical mirror image pairs.

was bypassed for this study\*. Eight mirror image variants of the identified base shape were used in the data collection tasks. These eight variants formed two tetrads of images. The first consisted of the base shape itself and the base shape rotated 180° laterally in depth, vertically in depth and 180° in-plane (Figure 5A). The second tetrad was created by performing the same transformations on the 90° rotation of the base shape (Figure 5B). The rationale for employing both tetrads was as follows. In selecting a base shape we might have

introduced a systematic bias whereby the axis of maximal symmetry tended to be close to vertical (as in an ‘M’) rather than close to horizontal (as in an ‘E’). If so, within tetrad 1 (e.g., Figure 5A), formed by mirror reflection of the upright base shape, the lateral mirror images would actually have been more similar to each other than the vertical mirror images, and would for that reason alone, quite independently of any tendency toward

\* This resulted in a high proportion of visual cells in the final analysis and is not indicative of the overall proportion of visual neurons in IT. Neurons that were bypassed could not be considered non-visual, as the set of visual stimuli was limited and may not have contained features that drove some neurons.

lateral mirror image confusion, have been expected to elicit more similar responses from IT neurons. However, within tetrad 2 (e.g., Figure 5B) formed by mirror reflection of the 90° rotated base shape, vertical mirror images should then have been more similar to each other and have elicited more similar responses. By including both tetrads, we ensured that any symmetry present in the base image would contribute equally to the similarity of stimuli forming lateral mirror image pairs and to the similarity of stimuli forming vertical mirror image pairs.

### *2.2.3 Behavioral Paradigm*

To determine whether neurons in IT respond more similarly to lateral mirror images than to vertical mirror images, eight mirror image variants of a single base shape were presented foveally on randomly interleaved trials to the fixating monkey. Events during a representative trial are shown in Figure 6. The monkey initially fixated a 0.6° blue spot at the center of the screen for 600 ms. Then one of the eight images was superimposed over the fixation spot for 600 ms. Following offset of the image, the monkey was required to fixate for an additional 600 ms. Continuous fixation throughout the trial was rewarded with a drop of juice. Any trial was terminated, and data from that trial discarded, if the monkey's gaze shifted outside an approximate 2° x 2° invisible window centered on the fixation spot. Neuronal data were collected until the monkey had successfully completed sixteen trials of each of the eight conditions. Data from a third monkey (monkey Ph) were collected while the monkey performed a slight variation of the task. Task parameters were identical to those for monkeys Op and Fi with two exceptions. One, the base shape was selected from a different set of stimuli (Figure 4B)

and two, the eight variants were randomly interleaved with images that the monkey had been exposed to in an orientation discrimination task (see Chapter 3). For the purposes of this study, only those eight conditions containing familiar but untrained images were considered. The reasons for including data from this monkey are twofold: (1) to determine whether effects observed in this study are specific to the stimuli used and (2) to extend the finding to a third monkey.

To determine whether the effects observed under foveal presentation persisted under peripheral presentation, data were collected while the eight mirror image variants were presented 4.8° to the right and left of the fixation. Timing in this task was identical to that of the foveal task. The eight images were presented at the two locations on separate, randomly interleaved trials. Data were collected until the monkey had successfully completed sixteen trials of each of the sixteen conditions (8 mirror image variants x 2 locations). Data were also collected from the same neurons during a block of the foveal task.

#### *2.2.4 Statistical Analysis*

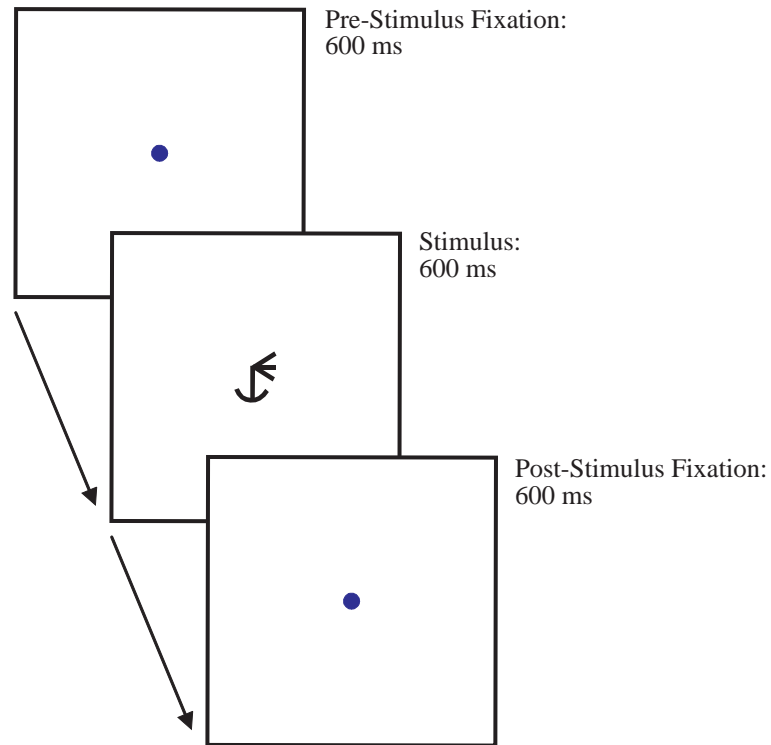
Trials were divided into two epochs for the analysis of neural activity. The “baseline” period was the 400 ms immediately preceding image onset. The “stimulus” period extended from 50-500 ms after image onset. The first 50 ms were excluded from analysis to account for the visual response latency of IT neurons. The length of the stimulus period was designed to include the period of strongest visual responsiveness. Neurons were initially assessed for visual responsiveness by comparing firing rates during the baseline and stimulus periods using a matched pairs *t*-test, evaluated at  $p <$



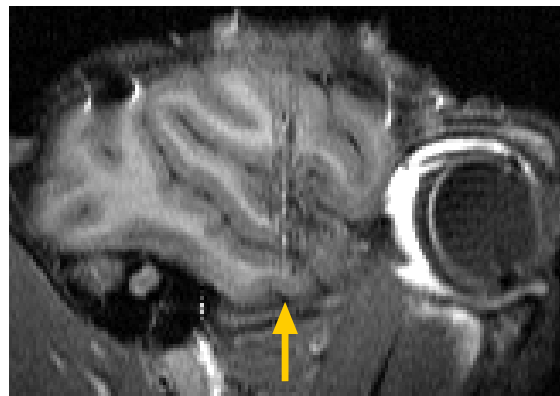
0.01. If this test was significant for at least one of the eight mirror image variants, then that neuron was included in the database for subsequent analysis.

To assess the degree to which neurons in IT exhibited lateral mirror image confusion, two methods of analysis were used. The first employed a series of *t*-tests to determine the counts of lateral and vertical mirror image pairs that evoked significantly different responses from the neurons. Specifically, for each pair of mirror images, for each neuron, a *t*-test was performed on the neural responses elicited by the two orientations during the stimulus period at a significance level of  $p < 0.05$ . This was done for each of the four lateral and four vertical mirror image pairs. Thus, a total of eight *t*-tests were performed on data from each neuron. The relative numbers of lateral and vertical mirror-image pairs evoking significantly different responses in the neurons were then compared using a Chi squared analysis.

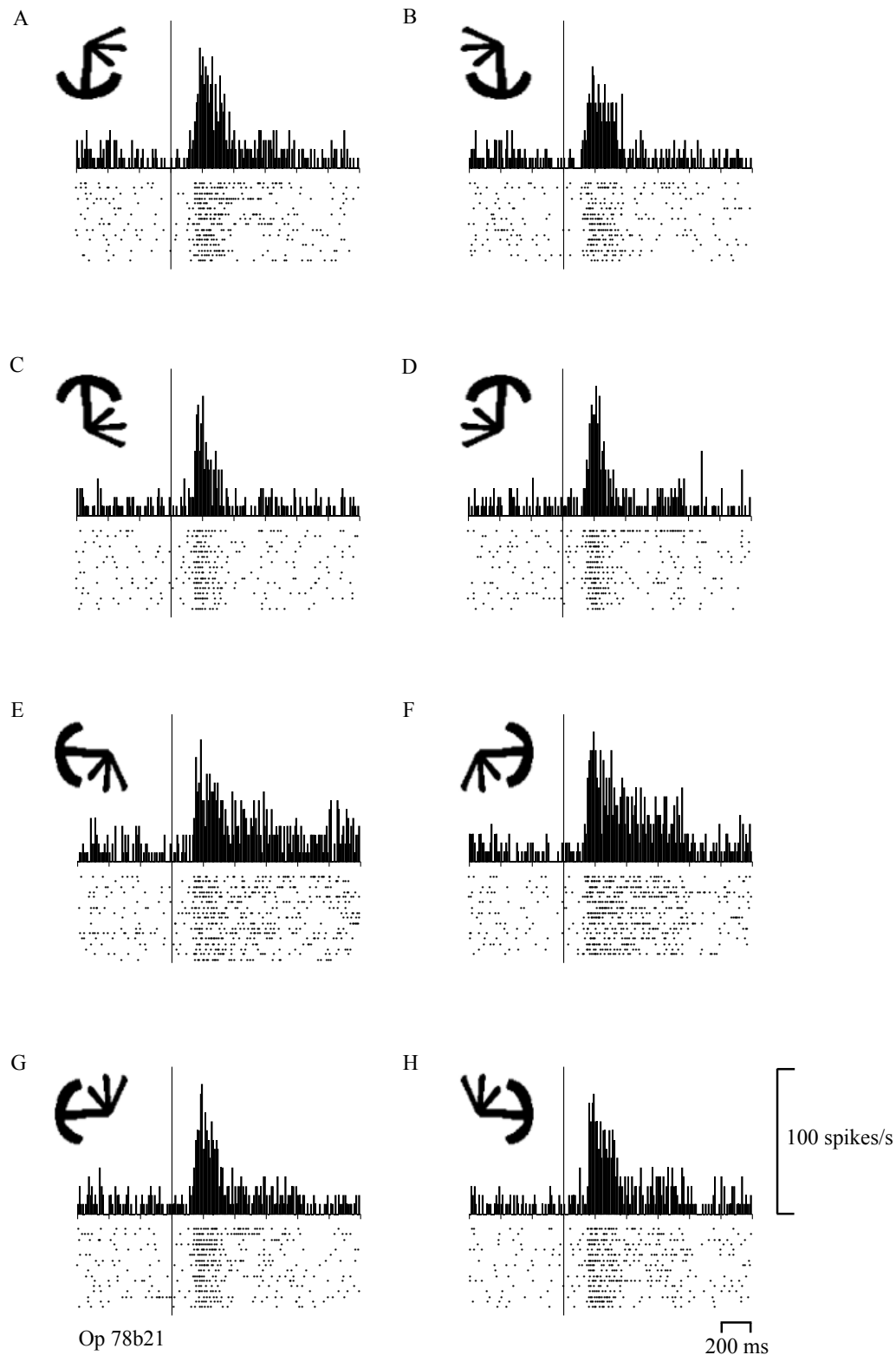
The second method of analysis was used to confirm the results obtained with the pair analysis. Specifically, neural responses during the stimulus period were analyzed using a two-factor ANOVA with lateral rotation and vertical rotation in depth as the two factors. One ANOVA was run on each tetrad, each evaluated at a criterion level of  $p < 0.05$ . Therefore, two ANOVAs were run on data collected from each neuron. The relative numbers of main effects of lateral rotation and vertical rotation across all neurons were compared using Chi squared analysis.



**Figure 6. Fixation task.** Each panel represents the screen in front of the monkey during successive epochs of a single representative trial. The trial began with the monkey fixating a central fixation spot for 600 ms. Then, one of eight mirror image variants was presented foveally or  $4.8^\circ$  to the left or right of fixation for 600 ms. The eight variants were presented separately on randomly interleaved trials.



**Figure 7. Parasagittal magnetic resonance image of the right hemisphere of monkey Op.** Arrow indicates the center of the recording zone. Also visible are guide-tube tracks in overlying tissue and a dark artifact from a titanium skull screw above the parieto-occipital cortex.



**Figure 8. Example of a single neuron exhibiting lateral mirror image confusion.** (A-H) Responses to eight orientations of the same shape presented at the fovea. An image was presented 600 ms after attainment of fixation and remained on for 600 ms. Data were aligned on the image onset (vertical line). Vertical calibration bar, 100 spikes/s; ticks on the horizontal axis, 200 ms; histogram bin width, 10 ms.

## 2.3 Results

### 2.3.1 Location of Recording Sites

Recording was carried out in anterior IT in the right hemisphere of two monkeys (monkeys Op and Fi) and in the left hemisphere of a third (monkey Ph). In all three animals, recording sites were lateral to the anterior medial temporal sulcus, and therefore were restricted to area TE. Recording in monkey Op was confined to frontal levels in the range anterior 18-22 mm as defined with respect to the interaural plane (Figure 7). The range of recording sites was 17-20 mm in monkey Fi, and 13-16 mm in monkey Ph. With respect to depth, recording sites in monkey Op were limited to the ventral aspect of the inferotemporal gyrus, whereas recording sites in monkeys Fi and Ph were localized to the lower bank of the superior temporal sulcus as well as the ventral aspect of the inferotemporal gyrus.

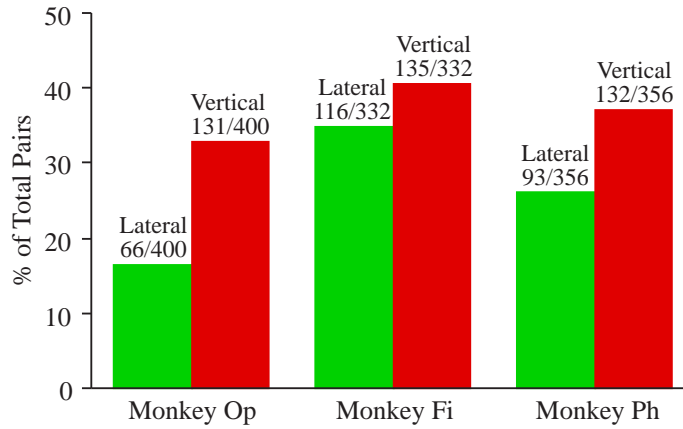
### 2.3.2 Foveal Presentation

Data were collected from 304 neurons during foveal presentation of mirror image stimuli (monkey Op:  $n = 111$ ; monkey Fi:  $n = 89$ ; monkey Ph:  $n = 104$ ). Of these neurons, 272 were visually responsive (monkey Op:  $n = 100$ ; monkey Fi:  $n = 83$ ; monkey Ph:  $n = 89$ ). Only these neurons were considered further. In examining the responses of neurons to presentation of mirror images, we observed clear instances in which neuronal responses to members of a lateral mirror image pair were more similar than responses to members of a vertical mirror image pair. One example neuron is shown in Figure 8. Histograms representing responses to lateral mirror images are beside each other (A-B,

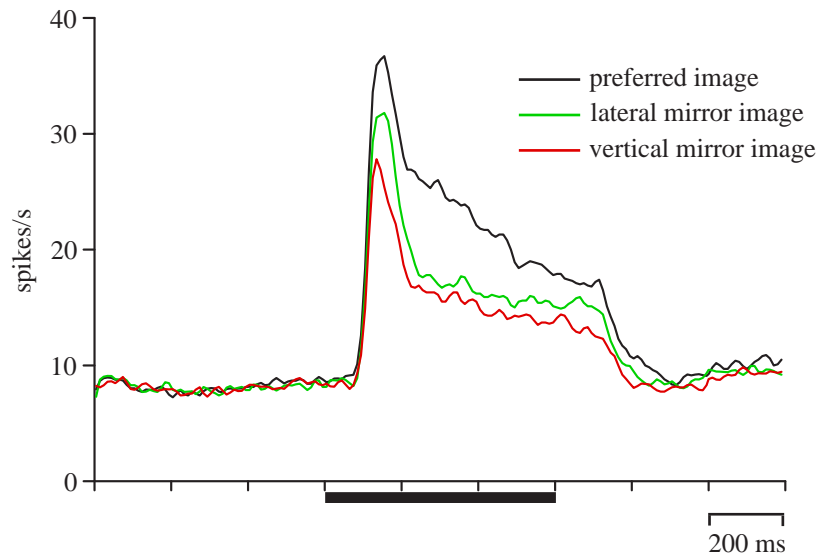
C-D, E-F, G-H) whereas histograms representing responses to vertical mirror images are juxtaposed vertically (A-C, B-D, E-G, F-H). Comparison of responses to members of mirror image pairs reveals that this neuron responded more similarly to members of lateral mirror image pairs than to members of vertical mirror image pairs.

In order to quantify this effect, we determined, for each neuron, the number of lateral and vertical mirror image pairs that evoked significantly different responses (see methods). The total numbers of lateral and vertical mirror image pairs meeting this criterion are shown separately for the three monkeys in Figure 9. For all three monkeys, instances of significant selectivity between vertical mirror images were more numerous than instances of significant selectivity between lateral mirror images pairs (Table 1). This effect was highly significant in monkey Op ( $X^2 = 28.45$ ,  $p < 10^{-7}$ ) and monkey Ph ( $X^2 = 9.88$ ,  $p < 0.002$ ), and presented as a non-significant trend in monkey Fi ( $X^2 = 2.31$ ,  $p > 0.1$ ).

We next asked to what extent individual neurons confused lateral mirror images more than vertical. That is, it might be the case that all neurons selective for pairs of images are more selective for vertical mirror image pairs than for lateral mirror image pairs. Alternatively, some neurons may be more selective for vertical pairs and some selective for more lateral pairs. To test this, we calculated the number of neurons that discriminated between more lateral mirror image pairs than vertical mirror image pairs ( $L > V$ ) and the number of neurons that discriminated between more vertical mirror image pairs than lateral mirror image pairs ( $V > L$ ). The counts of neurons are presented separately for each monkey in Table 1. While there existed neurons that discriminated between more lateral than vertical pairs, these neurons were outnumbered by those



**Figure 9. Lateral mirror image confusion for images presented at the fovea.** For all three monkeys, vertical mirror image pairs that elicited significantly different responses outnumbered lateral pairs that elicited significantly different responses. This effect was significant for monkeys Op ( $X^2 = 28.45$ ,  $p < 10^{-7}$ ) and Ph ( $X^2 = 9.88$ ,  $p < 0.002$ ), and approached significance for monkey Fi ( $X^2 = 2.31$ ,  $p > 0.1$ ).



**Figure 10. Time-course of lateral mirror image confusion.** Black line corresponds to the average population response to the preferred orientation, identified separately for each neuron. Green and red lines correspond to the average population responses to the lateral and vertical mirror image rotations of the preferred orientation, respectively. Curves are averages of responses from 272 neurons from monkeys Op, Fi and Ph. Data were aligned on the image onset. Black bar under horizontal axis represents duration of image presentation (600 ms). Tick marks on the horizontal axis are in increments of 200 ms; histogram bin width, 10 ms. Curves were smoothed according to the formula  $Y_n' = 0.25Y_{n-1} + 0.5Y_n + 0.25Y_{n+1}$ , where  $Y_n$  is instantaneous firing rate.

**Table 1. Significant differences in neuronal activity during the presentation of members of mirror-image pairs.** Image-pair summary is based on eight comparisons (four lateral and four vertical) from each neuron. In the neuronal summary, each cell was categorized according to the results of eight image-pair comparisons of its data. If activity differed significantly between more lateral pairs than vertical pairs, it contributed to the lateral > vertical count and vice versa.

Location	Subject	By image pair		By neuron	
		Lateral	Vertical	Lateral > Vertical	Vertical > Lateral
Fovea	Monkey Op	66/400	131/400	11/100	51/100
	Monkey Fi	116/332	135/332	22/83	31/83
	Monkey Ph	93/356	132/356	18/89	35/89
Ipsilateral hemifield	Monkey Op	21/132	55/132	3/33	22/33
	Monkey Fi	14/96	13/96	7/24	6/24
Contralateral hemifield	Monkey Op	24/132	60/132	4/33	19/33
	Monkey Fi	23/96	38/96	4/24	16/24

**Table 2. Lateral mirror image confusion analysis by ANOVA.** Factors were lateral mirror image rotation and vertical mirror image rotation. Analysis was done separately for each of the two tetrads, leading the number of possible main effects to be twice the number of neurons studied.

Location	Monkey	Number of neurons	Main effect of lateral rotation only	Main effect of vertical rotation only	Interaction only	2 main effects	Main effect of lateral rotation + interaction	Main effect of vertical rotation + interaction	2 main effects + interaction	No significant effects	Total main effects of lateral rotation	Total main effects of vertical rotation
Fovea	Monkey Op	100	17	52	7	18	4	7	6	89	45	83
	Monkey Fi	83	23	30	12	19	6	11	25	40	73	85
	Monkey Ph	89	19	48	11	19	3	6	11	61	52	84
Ipsilateral Hemifield	Monkey Op	33	3	25	3	2	0	4	5	24	10	36
	Monkey Fi	24	7	5	2	1	0	2	1	30	9	9
Contralateral Hemifield	Monkey Op	33	4	23	5	6	1	5	3	19	14	37
	Monkey Fi	24	5	11	5	5	0	2	6	14	16	24



discriminating between more vertical than lateral pairs in all three monkeys. This effect was significant in two of the three monkeys (monkey Op:  $X^2 = 25.81$ ,  $p < 10^{-6}$ ; monkey Fi:  $X^2 = 1.53$ ,  $p > 0.2$ ; monkey Ph:  $X^2 = 5.45$ ,  $p < 0.02$ ).

It could be argued that the differences in firing rates observed in individual pairs were due simply to variance in the cells' responses and do not reflect systematic differences. The pair-wise comparisons do not take into account the variance of the cells across all conditions. Therefore, to confirm the results of the pair analysis, we analyzed the responses to lateral and vertical mirror images using an ANOVA (see methods). The results of this analysis are in accord with the previous analyses in that for all three monkeys, there were a greater number of main effects of vertical rotation than main effects of lateral rotation (Table 2). This effect was highly significant in two of the three monkeys (monkey Op:  $X^2 = 16.59$ ,  $p < 0.00005$ ; monkey Fi:  $X^2 = 1.74$ ,  $p > 0.1$ ; monkey Ph:  $X^2 = 12.18$ ,  $p < 0.0005$ ).

### *2.3.3 Time-course of Lateral Mirror Image Confusion*

The analyses discussed thus far considered response rates calculated over a long segment of the image presentation time. We next asked whether the lateral mirror image confusion expressed in IT neurons varied as a function of time. If this effect was determined by properties of feed-forward pathways, or intra-areal circuits, then it might be expected that the difference in selectivity for lateral and vertical mirror images would arise early in the visual response. If the effect was a consequence of feedback from areas of higher order or of behavioral responses to the stimuli, then it might be expected that the difference would arise later in the response. To test this, population histograms were

created from data from the 272 visual neurons recorded from all three monkeys. The preferred image of each neuron was identified by determining which of the eight mirror image variants elicited the highest firing rate during the stimulus period. The responses to the preferred image were averaged across all 272 neurons for each 10 ms bin and shown as a function of time in Figure 10 (black line). The lateral mirror image and vertical mirror image of the preferred image of each neuron were identified and the responses to each were averaged across the population of neurons. The population histograms for the lateral and vertical mirror images are presented in Figure 10, in green and red, respectively. It is clear that the responses elicited by lateral mirror images (the black and green lines) are more similar than the responses elicited by vertical mirror images (the black and red lines). While this effect appears to be slightly stronger in the initial portion of the response, it is sustained throughout the entire visual response. The fact that the difference in responses to lateral and vertical mirror images emerges very early, within 20-30 ms of the visual response onset, suggests that the effect is not a result of any behavioral response, such as eye movements\*. This finding suggests that lateral mirror image confusion may result from either feed-forward or intra-areal connections.

#### *2.3.4 Peripheral Presentation*

To determine whether neurons exhibited lateral mirror image confusion for images presented in the periphery, data were collected from 57 neurons (monkey Op:  $n =$

\* Analysis of eye movement data collected during foveal presentation revealed that the amplitude of the eye movements never exceeded  $0.5^\circ$ , and movements were highly stereotyped in each monkey, not varying systematically across either shape or mirror image variant. Deviations in eye position began around 140 ms after stimulus onset although they did not achieve full amplitude until several tens of milliseconds later. With a visual response latency of 80-100 ms, the earliest impact of this displacement on neuronal activity in IT would have occurred 220 ms after initial onset of the stimulus and therefore long after the advent of signals differentiating between lateral and vertical mirror images.

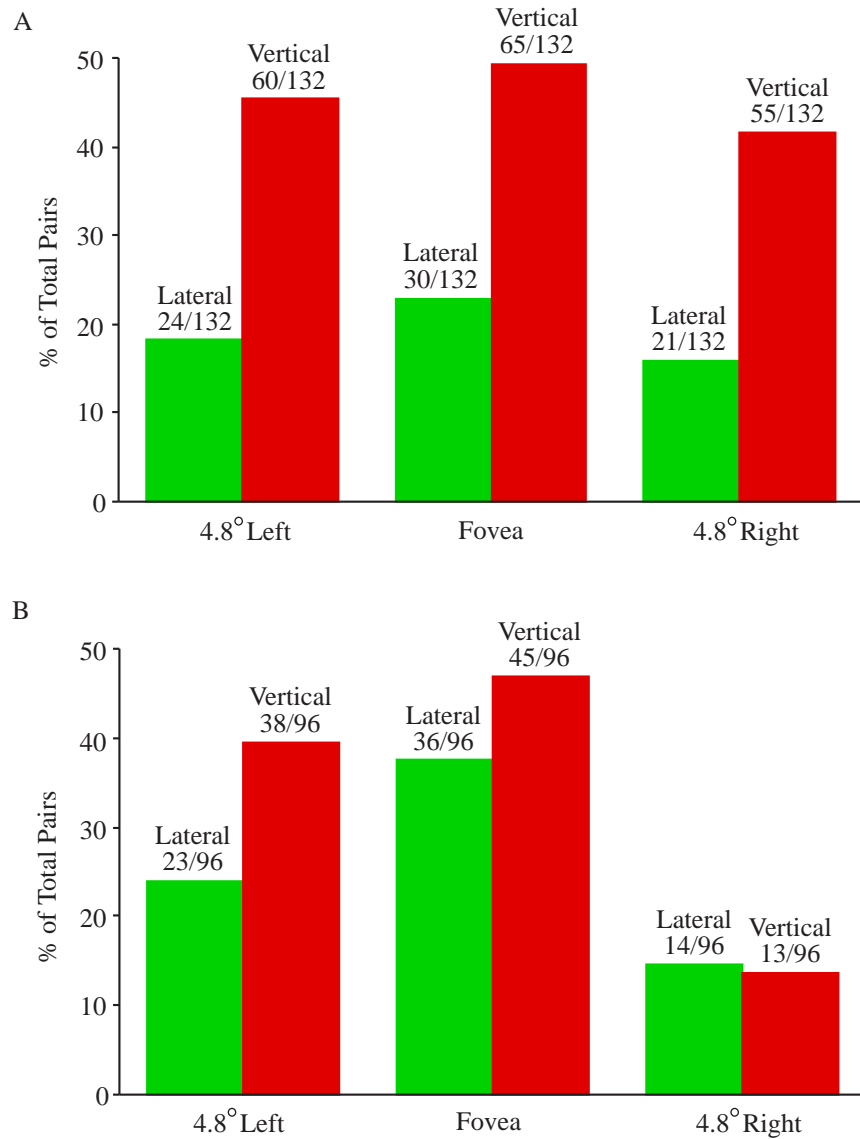
33; monkey Fi:  $n = 24$ ), monitored during presentation of images at the fovea and  $4.8^\circ$  to the right and left of fixation. We determined, for each neuron, the numbers of lateral and vertical mirror image pairs that evoked significantly different responses at each of the three locations. The total numbers of lateral and vertical mirror image pairs meeting this criterion at each of the three locations are presented separately for the two monkeys in Figure 11 and in Table 1. In monkey Op, across all three locations, instances of significant selectivity between vertical mirror images were more numerous than instances of significant selectivity between lateral mirror images pairs. This effect was highly significant for all three locations (fovea:  $X^2 = 20.14$ ,  $p < 10^{-5}$ ; contralateral:  $X^2 = 22.63$ ,  $p < 10^{-5}$ ; ipsilateral:  $X^2 = 21.36$ ,  $p < 10^{-5}$ ). A loglinear analysis revealed a significant effect of mirror reflection (lateral or vertical;  $X^2 = 60.87$ ,  $p < 0.0001$ ), but no effect of location (fovea, left hemifield or right hemifield;  $X^2 = 3.53$ ,  $p > 0.1$ ) and no interaction between location and mirror reflection ( $X^2 = 0.15$ ,  $p > 0.9$ ). In monkey Fi, the same pattern of selectivity was observed for images presented at the fovea ( $X^2 = 1.7297$ ,  $p > 0.1$ ) and the contralateral hemifield ( $X^2 = 5.4061$ ,  $p < 0.05$ ), but not the ipsilateral hemifield ( $X^2 = 0.0431$ ,  $p > 0.8$ ). A loglinear analysis revealed a non-significant trend toward a main effect of mirror reflection ( $X^2 = 2.96$ ,  $p > 0.08$ ), a main effect of location ( $X^2 = 34.45$ ,  $p < 0.0001$ ) but no interaction between mirror reflection and location ( $X^2 = 2.46$ ,  $p > 0.2$ ).

We found a similar result when considering the selectivity of individual neurons. That is, we found that there were a greater number of neurons that were more selective for vertical rotations ( $V > L$ ) than neurons that were more selective for lateral rotations ( $L > V$ ), at both the fovea (monkey Op:  $X^2 = 10.70$ ,  $p < 0.002$ ; monkey Fi:  $X^2 = 2.88$ ,  $p > 0.1$ ) and in the contralateral hemifield (monkey Op:  $X^2 = 9.78$ ,  $p < 0.002$ ; monkey Fi:  $X^2 =$

7.2,  $p < 0.01$ ). In the ipsilateral hemifield, the same pattern was observed in monkey Op ( $X^2 = 14.44$ ,  $p < 0.0002$ ), but not in monkey Fi ( $X^2 = 0.08$ ,  $p > 0.7$ ). These results indicate that, for the most part, the tendency for IT neurons to confuse lateral mirror images is independent of visual field location within the tested range.

To confirm these findings, we analyzed the responses to images at each of the three locations by an ANOVA (see methods). This analysis revealed a similar result to that of the pair analysis (Table 2). For all three locations in monkey Op, there were a greater number of main effects of vertical rotation than main effects of lateral rotation, and at all three locations this effect was highly significant (fovea:  $X^2 = 10.05$ ,  $p < 0.002$ ; contralateral:  $X^2 = 16.90$ ,  $p < 0.00005$ ; ipsilateral:  $X^2 = 22.56$ ,  $p < 0.00001$ ). In monkey Fi, while the same pattern was observed in the contralateral hemifield, it did not reach significance ( $X^2 = 2.74$ ,  $p > 0.09$ ). At the other two locations, there was no effect (fovea:  $X^2 = 0.68$ ,  $p > 0.6$ ; ipsilateral:  $X^2 = 0$ ,  $p = 1$ ). The absence of an effect for this monkey in the ipsilateral hemisphere may be due to several factors. In general, neurons exhibiting visual responses were more difficult to find, and visual responses were weaker, than in monkey Op. Although recording sites appeared to be comparable, as determined from the MRI images, it is possible that there were subtle differences in recording locations between the two monkeys. In addition, receptive fields were not mapped in the present study. It might be the case that receptive field properties varied systematically between the two monkeys such that ipsilateral responses in monkey Fi were more variable and less selective.

Finally, analysis of eye movement data during peripheral presentation revealed that for neither monkey was there a deflection of eye movements greater than  $0.5^\circ$  and



**Figure 11. Lateral mirror image confusion in neurons for images presented at the fovea and in the periphery.** Data is shown separately for monkeys Op (A) and Fi (B). For monkey Op, vertical mirror image pairs that elicited significantly different responses outnumbered lateral pairs that elicited significantly different responses for images presented at the fovea, in the contralateral (left) hemifield and in the ipsilateral (right) hemifield. This effect was significant in monkey Op for images presented at all three locations (left hemifield:  $X^2 = 22.63$ ,  $p < 10^{-5}$ ; fovea:  $X^2 = 20.14$ ,  $p < 10^{-5}$ ; right hemifield:  $X^2 = 21.36$ ,  $p < 10^{-5}$ ). The same pattern was observed for monkey Fi for images presented at the fovea and in the contralateral (left) hemifield. This effect was significant in monkey Fi for images presented in the contralateral (left) hemifield, but not at the fovea (left hemifield:  $X^2 = 5.41$ ,  $p < 0.05$ ; fovea:  $X^2 = 1.73$ ,  $p > 0.1$ ; right hemifield:  $X^2 = 0.04$ ,  $p > 0.8$ ).

that eye position did not vary systematically across conditions. This supports the idea that the lateral mirror image effect observed at the fovea and in the periphery does not reflect any behavioral response of the monkey.

### *2.3.5 Mirror Image Preference in the Two Hemifields*

If lateral mirror image confusion arises from connectivity between neurons in opposite hemispheres preferring opposite members of a lateral mirror image pair, then it might be expected that neurons would prefer opposite members of a lateral mirror image pair presented in the two hemifields. This follows from the finding that information in the ipsilateral portion of the receptive field of IT neurons is accessed via input from neurons in the opposite hemisphere (Gross et al., 1977). To test whether neurons prefer opposite members of lateral mirror image pairs in the two hemifields, we analyzed the stimulus preferences of neurons recorded from during the presentation of mirror image stimuli in the left (contralateral) and right (ipsilateral) hemifields. In all cases where a neuron significantly discriminated between members of a mirror-image pair in both hemifields, we asked whether the pattern of preference was the same or reversed across hemifields. In 34 out of 39 cases involving a vertical pair, the same member was preferred in the two hemifields. In 3 out of 4 cases involving a lateral pair, different members were preferred. The low frequency of lateral pairs that neurons were selective for in both hemifields prevented us from carrying out a statistical analysis of these data. Therefore, while there are greater cases where the preference for lateral mirror images was reversed in the two hemifields, further recording would be necessary to conclude if this was a statistically significant finding.

### 2.3.6 Control for Symmetry

In including the second tetrad of mirror images, we attempted to circumvent the possibility that inherent symmetry alone would result in the effect observed (see methods for rationale). However, it might be the case that neurons fire more strongly or more selectively for the tetrad that contained images with more symmetry about their vertical axes, thus leading to a decrease in selective responses for lateral compared to vertical mirror images. To test this possibility, we first asked whether there was a tendency for one tetrad to produce greater confusion than the other. For each neuron, we computed a lateral mirror image confusion index for each of the two tetrads within the testing octet. The index was calculated as  $(l - v) / (l + v)$ , where  $l$  is the total number of lateral pairs for a given shape that evoked significantly different responses and  $v$  is the total number of vertical pairs for the same shape that evoked significantly different responses. Across all neurons tested with a given shape, we then counted cases in which the first tetrad elicited greater confusion and cases in which the second tetrad elicited greater confusion. On average (across all 25 shapes), the tetrad with the greater number of counts exceeded the tetrad with the lesser number of counts by a factor of 2.9. The significance of these counts was evaluated using Monte Carlo simulations to determine the distributions of tetrads that would be expected by chance. This analysis revealed that the ratio of tetrads was significantly greater than the ratio expected by chance ( $p < 0.01$ ). Thus, for some or all shapes, one tetrad must have tended to produce greater confusion than the other. However, when we asked whether, across all 25 shapes, the tetrad yielding greater confusion tended to be the one eliciting stronger activity, we found that the correlation

was not significant ( $r = -0.1549$ ,  $p > 0.4$ ). Likewise, the tetrad yielding greater confusion did not tend to be the one eliciting greater selectivity ( $r = -0.0316$ ,  $p > 0.8$ ).

## 2.4 Discussion

We recorded single-neuron activity in the inferotemporal cortex of three monkeys during passive fixation of mirror image stimuli. We found that neurons respond more similarly to presentation of lateral mirror images than to presentation of vertical mirror images, and that this effect occurs for both foveal and peripheral presentation. We observed that signals discriminating lateral and vertical mirror images arise early in the visual response, suggesting that feed-forward or intra-areal connectivity may give rise to the observed effect.

This is the first demonstration that neurons in any brain area respond more similarly to lateral mirror images than to vertical mirror images. Previous studies have provided only incidental examples of lateral mirror images evoking similar response strengths in IT neurons. In one study, Tanaka et al. (1991) presented orientation tuning curves for eight IT neurons responsive to bilaterally symmetric stimuli. In the case of any bilaterally symmetric image, a  $180^\circ$  rotation in the viewing plane is equivalent to a mirror transformation about some axis (e.g., an “M” rotated  $180^\circ$  in-plane produces the vertical mirror image of an “M”). In all but one of the eight neurons, the difference in firing rate was greater for vertical than for lateral mirror image pairs. Although interpreted by the authors strictly in terms of orientation selectivity, these results are in accord with our findings. Two other studies have provided examples of neurons



exhibiting similar firing rates for lateral mirror images, but did not include the necessary comparison to vertical mirror image selectivity. In the first study, Perrett et al. (1991) examined the responses of STS neurons to faces rotated laterally in depth. For such bilaterally symmetric stimuli, certain rotations in depth form mirror images (e.g., right and left profiles of faces). The authors found that there exists a small subset of neurons that exhibited bimodal tuning curves such that they responded similarly to lateral mirror images. The second study examined responses of IT neurons to wire objects rotated in depth (Logothetis et al. 1995). These objects were sufficiently non-self-occluding to form pseudo-mirror images at certain orientations. As with the face selective neurons described above, these authors identified instances in which neurons exhibited bimodal tuning, preferring both lateral mirror images above all other orientations. While these examples do not provide conclusive evidence for lateral mirror image confusion in IT, they are consistent with the results of the present study.

The finding that lateral mirror images evoke more similar responses in IT neurons than do vertical mirror images parallels, and might provide the neural basis for, the increase in perceived similarity of lateral compared to vertical mirror images demonstrated behaviorally. While the degree to which the monkeys in the present study perceived the mirror images as similar was not measured (i.e. the degree to which monkeys confused mirror images), previous behavioral studies have shown that monkeys do exhibit lateral mirror image confusion (Riopelle et al., 1964; Hamilton and Tieman, 1973). This behavioral confusion has also been demonstrated in human primates (Bornstein et al., 1978; Huttenlocher, 1967; Serpell, 1971; Rudel and Teuber, 1963;

Sekuler and Houlihan, 1968), as well as octopus (Sutherland, 1960) and pigeons (Todorin and Blough, 1983).

The idea that neural processes within IT subserve the perceptual confusion of complex visual stimuli is supported by findings in two recent neurophysiological studies. Miyashita et al. (1993) compared the selectivity of IT neurons for fractal patterns to the perceptual similarity of those patterns, as rated by human observers. They found that neurons tended to fire more similarly to pairs of patterns that were rated as highly similar by humans. In a more recent study, Op de Beeck et al. (2001) reported that the responses of IT neurons to parameterized stimuli were highly congruent with the degree to which the monkeys confused those same stimuli. These studies support the findings of the experiments described here in suggesting that the neural underpinnings of perceived similarity may reside in IT.

The results of the present study are congruent with former theories of lateral mirror image confusion. One theory concerns the anatomical derivation of this phenomenon. Specifically it suggests that neurons in one hemisphere activated by a particular image are connected with neurons in the opposite hemisphere activated by the lateral reflection of that image (Corballis and Beale, 1970). It has been demonstrated that connections exist between the inferotemporal cortices of the two hemispheres through the corpus callosum (Gross et al., 1977; Desimone et al., 1980). Moreover, information in the ipsilateral portion of IT receptive fields is received via callosal connections from the opposite hemisphere (Gross et al., 1977). We tested this theory by considering whether preference for members of lateral mirror image pairs is reversed in the two hemifields. We found that there was a trend in this direction. However, there were too few cases in

which IT neurons discriminated between lateral mirror images in both hemifields to allow establishing the significance of this trend. The fact that we found so few significant pairs is consistent with our general finding that neurons are poor at discriminating lateral mirror images.

The second theory suggests that the relative difficulty in discriminating lateral mirror images is not a failure of the visual system, but rather an adaptive phenomenon (Gross and Bornstein, 1978). Given that lateral reversals observed in nature almost always result from changes in viewpoint, they offer little information for the purposes of recognition. Thus, a system dedicated to recognition and perception would not need to encode such reflections. In contrast, vertical reversals rarely come about from changes in viewpoint and therefore do convey important information about the object. Our finding that neurons in IT are less able to discriminate lateral mirror images than vertical mirror images, is consistent with this theory. It is unclear, however, whether this dichotomy arises during development, or is an evolutionary adaptation.

## **2.5 Summary**

In conclusion, we have demonstrated that neurons in IT discriminate lateral mirror images less effectively than vertical mirror images. This phenomenon may provide the neural basis for lateral mirror image confusion observed in behavior and suggests that the activity of IT neurons may reflect the perceived similarity of complex objects.

## **Chapter 3**

### **Effects of Training to Discriminate Shape Orientation on Neural Responses in Macaque Inferotemporal Cortex**

#### **3.1 Introduction**

As we gain greater visual experience with a particular object, the manner in which we perceive that object changes (Sheinberg and Logothetis, 2001). It follows, then, that the ability to recognize and discriminate complex objects depends on visual experience. Recent research has been aimed at identifying the site of such experience-based changes in the brain. A widely held view is that the increased ability to discriminate complex objects is due to changes in neuronal responses for those objects, and that these changes take place in the inferotemporal cortex (Wallis and Bulthoff, 1999; Hasegawa and Miyashita, 2002; Sheinberg and Logothetis, 2001). However, studies examining effects of training on IT neurons have produced contradictory results (Miyashita et al., 1993; Logothetis and Pauls 1995; Vogels and Orban, 1994b; Sakai and Miyashita, 1994; Logothetis et al., 1995; Kobatake et al., 1998; Erickson et al., 2000; Messinger et al., 2001).

First, it is unclear whether discrimination training affects the overall response strength of neurons for trained objects. Several studies have shown that the response strengths evoked by trained stimuli are greater than those evoked by untrained stimuli (Miyashita et al., 1993; Sakai and Miyashita, 1994; Kobatake et al., 1998). However, in one case responses evoked by trained and untrained stimuli were recorded in separate

animals (Kobatake et al., 1998). In two other cases, neurons were chosen based on their responsiveness to trained stimuli only (Sakai and Miyashita, 1994; Miyashita et al., 1993). These factors call into question the reliability of these results.

Second, while it has been hypothesized that discrimination training leads to increases in the selectivity of IT neurons for trained stimuli, this hypothesis has been supported by only a subset of training studies (Logothetis et al., 1995; Logothetis and Pauls, 1995; Kobatake et al., 1998). Studies by Erickson et al. (2000) and Vogels and Orban (1994b) failed to find significant changes in the selectivity of neurons in IT. However, in the former case, training lasted for only one day. In the latter case, monkeys were trained to discriminate line gratings, stimuli which do not typically evoked strong selective responses in IT neurons (Tanaka et al., 1991).

Third, it is not clear how widespread these proposed changes in neuronal selectivity are across neurons in IT. Logothetis and colleagues (Logothetis and Pauls, 1995; Logothetis et al., 1995) claimed that training induces strong changes in selectivity in a small subset of neurons. However, the authors did not carry out any statistical analyses or population measures to support this. In contrast, Kobatake et al. (1998) suggests that changes in selectivity occur as subtle increases in selectivity for trained stimuli across large numbers of neurons.

These three issues were recently addressed by Baker et al. (2002) in a study examining the effects of training monkeys to perform a feature conjunction task on neuronal responses in IT. In particular, they trained monkeys to discriminate between “baton” stimuli, which necessitated learning conjunctions of top and bottom features. They found that training on this task resulted in a subtle increase in selectivity for trained

batons across the population of neurons studied, but had no significant effect on the response strength of neurons to the preferred batons.

The study by Baker et al. (2002), as well as most of the previous studies discussed above, required monkeys to discriminate between stimuli containing different features. It is unclear, however, whether responses in IT can be modulated by training to discriminate stimuli that differ only in orientation while containing the same features. Therefore, in the present study, we investigated the effects of training on the responses of IT neurons when monkeys were required to discriminate between stimuli that differed in orientation. One other study has examined the effects of orientation discrimination training on the responses of IT neurons. Vogels and Orban (1994b) trained monkeys to perform a successive orientation discrimination task with line gratings and then compared the selectivity of IT neurons for trained and untrained orientations. The authors failed to find a significant effect of training on neuronal selectivity for orientation. However, the stimuli used in this study were simple line gratings which may not have sufficiently driven IT neurons, as IT neurons, particularly in more anterior regions, typically respond best to more complex stimuli (Tanaka et al., 1991). Thus, it remains unclear whether shape orientation discrimination training with appropriate stimuli would result in changes in neuronal responses in IT. Therefore, in this study, we set out to determine whether the response strength and selectivity of neurons in IT can be modulated by orientation discrimination training.

In Chapter 2, we demonstrated that, in certain cases, the orientation selectivity of IT neurons is anisotropic with respect to the relative orientation of images. Specifically, we found that neurons in IT respond with more similar response strengths to shapes that

are lateral mirror images of each other than to shapes that are vertical mirror images of each other. It is unclear whether this effect is hard wired in the brain or whether it is modifiable by experience. Therefore, in the present study, we ask whether training monkeys to discriminate lateral and vertical mirror images alters the degree to which IT neurons exhibit lateral mirror image confusion.

To better understand the functional implications of neuronal activity, it is particularly important to determine how this activity relates to the perception of the monkey. It has been demonstrated by several groups that the selectivity of IT neurons for complex shapes parallels the ability of monkeys to associate and discriminate those same shapes (Messinger et al., 2001; Op de Beeck et al., 2001). It is unclear, however, to what degree orientation selective responses in IT reflect the perception of shape orientation. If the perception of shape orientation is encoded in the responses of IT neurons, then it might be expected that the selectivity of neurons for orientation would parallel the monkey's ability to discriminate those orientations. In this study, we also investigate whether the ability of neurons to discriminate shape orientation correlates with the ability of the monkey to discriminate shape orientation.

Finally, it is unclear how task context affects the responses of IT neurons for complex shapes. It might be the case that the response properties of neurons differ depending on whether the monkey is passively fixating shapes, or actively processing them for the purposes of performing a task. Although several groups have failed to find an effect of task context on the visual responses of IT neurons (Baker et al., 2002; Lehky and Tanaka, 2002), it is not clear whether this would be the case for orientation selective responses in IT as well. In this study we ask whether the strength and selectivity of

neuronal responses for shape orientations differ depending on whether the monkey is passively viewing orientations or actively discriminating them.

In the present study, we investigated the effects of shape orientation discrimination training on the responses of IT neurons. We posed the following specific questions. 1) Does training affect the overall response strength of neurons for trained vs. untrained orientations? 2) Does training increase the selectivity of neurons for trained orientations? 3) Does any change in orientation selectivity as a result of training occur across many neurons in the population or only in a small subset of neurons? 4) Does training to discriminate between both lateral and vertical mirror image pairs change the degree to which neurons exhibit lateral mirror image confusion? 5) Does the neuronal selectivity for shape orientations presented in the DMS task correlate with the monkeys' ability to discriminate those orientations? 6) Does neuronal selectivity for trained orientations depend on task context? To address these questions, we trained monkeys to discriminate among four mirror reflected orientations of each of ten shapes in the context of a delayed match to sample (DMS) task. We then recorded responses of single IT neurons while monkeys performed a passive fixation task in which the following stimuli were presented: (1) trained orientations of trained shapes, (2) untrained orientations of untrained shapes and (3) untrained orientations of trained shapes. By including both (2) and (3), we were able to investigate the specificity of training-induced effects both for particular shapes and for particular orientations of those shapes. We also recorded from neurons while monkeys performed the DMS task using only those orientations on which they had been trained. This allowed us to ask whether neuronal selectivity for trained



orientations correlated with the monkey's ability to discriminate those orientations as well as whether it was affected by task context (active discrimination vs. fixation).

We found that while training did not affect the overall response strength of neurons, it did enhance neuronal selectivity for trained orientations. This effect occurred as a subtle increase in selectivity across the entire population of neurons studied, rather than a strong effect in a small subset of neurons. The increase in selectivity occurred for lateral and vertical mirror image pairs alike, so that the relative selectivity remained better for vertical than for lateral mirror image pairs. We found that there was a moderate correlation between the selectivity of neurons for shape orientation and the ability of the animals to discriminate those orientations. Finally, we observed that neuronal selectivity for trained orientations did not depend on task context.

## 3.2 Methods

### 3.2.1 Subjects

Two adult male rhesus monkeys (*Macaca mulatta*) weighing 7.5-9.0 kg were used in this experiment. Their laboratory designations were Ph and Op. General surgical and training procedures are described in Appendix A. Training specific to this experiment is described below.

### 3.2.2 Stimuli

The stimuli were twenty white, asymmetric shapes, each subtending approximately 3° of visual angle in height and width (Figure 12A-B). One set of ten objects was used in training for monkey Ph (Figure 12A) and the other in training for monkey Op (Figure

12B). Shape was counterbalanced against training status across the two monkeys in the sense that the shapes used in training with monkey Ph were used as untrained controls in monkey Op and vice versa. Untrained shapes were presented in a fixation task prior to data collection and thus were familiar to the monkey, but had never been seen in the context of a discrimination task.

### *3.2.3 Training*

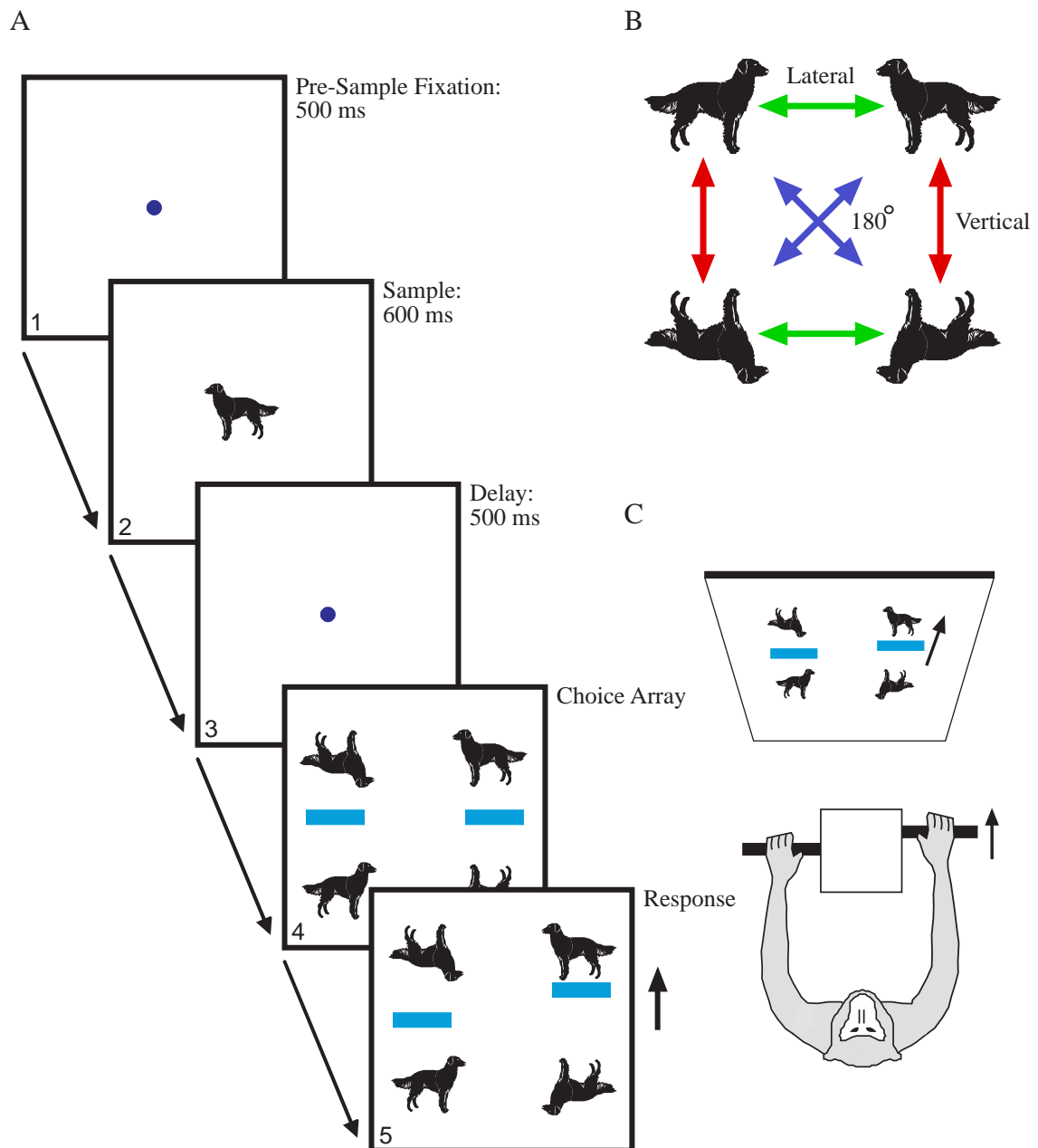
Monkeys were initially trained for several months to discriminate among four orientations of each of the ten base shapes within the confines of a delayed match-to-sample task (DMS). The four orientations were formed by rotating the base shape 180° laterally in depth, 180° vertically in depth and 180° in-plane (Figure 13B). They formed two pairs of lateral mirror images (green arrows), two pairs of vertical mirror images (red arrows) and two pairs of 180° rotations (blue arrows). On each trial, the monkey was presented with one orientation of the shape and was required, after a delay, to choose the same orientation from among all four (Figure 13A). Events occurring sequentially during a representative trial are presented in Figure 13A. The trial began with the monkey fixating a 0.6° blue fixation spot for 500 ms (panel 1). Then, a sample stimulus was superimposed over the fixation spot for 600 ms (panel 2). The sample was then extinguished and the monkey was required to maintain fixation during a 500 ms delay period (panel 3). Failure to maintain fixation within a 2° x 2° window throughout the fixation, sample and delay periods, resulted in the termination of the trial and no juice reward. After the delay period, an array of the four orientations was presented (Figure 13A, panel 4), at which time the monkey was no longer required to maintain fixation.



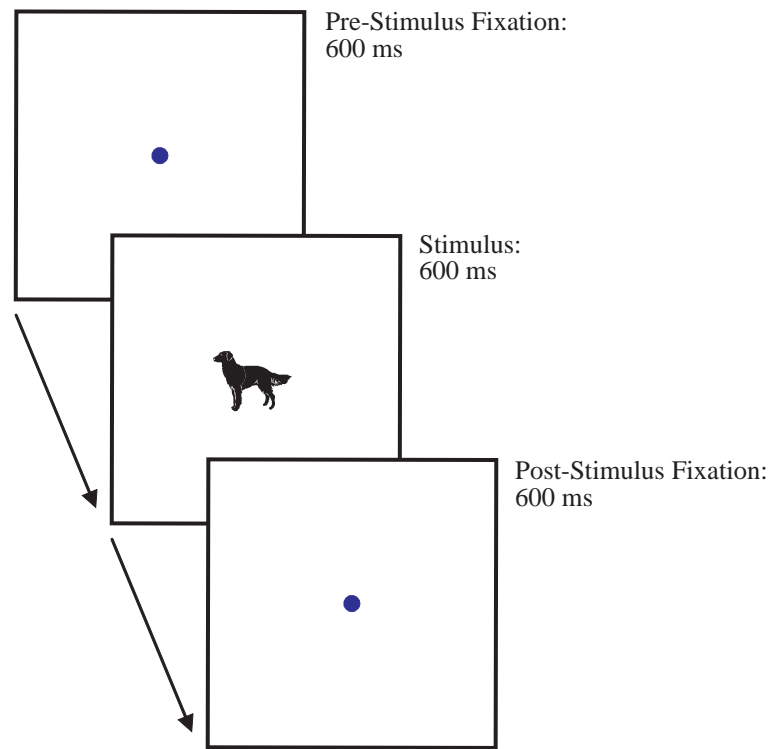
**Figure 12. Stimuli used in training and data collection.** Stimuli were white, chiral shapes approximately  $3^\circ$  in height and width, presented against a black background. A. Shapes used in orientation discrimination training for monkey Ph and used as untrained controls in monkey Op. B. Shapes used in orientation discrimination training for monkey Op and used as untrained controls in monkey Ph.

The arrangement of the four orientations was varied randomly among the four locations from trial to trial. The monkey was required to choose the shape with the same orientation as the sample, making his choice by moving one of two levers affixed to the primate chair either forward or backward (Figure 13C). For example, to choose the upper right probe, the monkey pushed the right lever forward. A rectangle on the display, acting as the lever's proxy, then moved upward, reflecting the monkey's choice (Figure 13A, panel 5). Monkeys had 3 seconds

from probe onset to respond. Correct choices were rewarded with a drop of juice. On incorrect trials a time penalty was imposed, during which the matching probe flashed to indicate the correct choice and no juice reward was given. The four orientations were presented as samples with equal frequency across all trials. The four trial conditions were interleaved in pseudorandom order subject to the constraint that the monkey had to complete successfully two trials of each type in each block of eight successful trials. Over several months, the monkey was trained on the ten different shapes (ten sets of four orientations), until his performance was above 65% on each (chance performance = 25%).



**Figure 13. Delayed match to sample (DMS) task used in training and data collection.** A. Each panel represents the screen in front of the monkey during successive epochs of a single representative trial. While the monkey fixated a central spot, a sample, one of the four orientations of the shape, was presented foveally (600 ms). Following a delay of 500 ms, the fixation spot was extinguished and the four orientations of the shape were presented simultaneously. The monkey was required to choose the shape with the same orientation as the sample. The corresponding rectangle moved toward the chosen probe to indicate a response. On error trials, the correct probe was flashed on and off to indicate the correct choice. B. Four orientations used in discrimination training formed by mirror inversion and rotation of the base shape. These four orientations formed two lateral mirror image pairs (green arrows), two vertical mirror image pairs (red arrows) and two 180° rotation pairs (blue arrows). C. The monkey indicated his choice by moving one of two levers affixed to the primate chair either forward or backward.



**Figure 14. Fixation task used in data collection.** Each panel represents the screen in front of the monkey during successive epochs of a single representative trial. The trial began with the monkey fixating a central fixation spot for 600 ms. Next, one of eight orientations of a trained shape or one of eight orientations of an untrained shape was presented foveally for 600 ms. Following offset of the shape, the monkey was required to maintain fixation for an additional 600 ms. Monkeys were rewarded with a drop of juice for maintaining fixation throughout the trial. In monkey Ph, orientations of trained and untrained shapes were randomly interleaved. In monkey Op, orientations of trained and untrained shapes were presented in separate blocks, and within each block, the eight orientations were randomly interleaved. The order in which these two blocks were run during data collection was varied from session to session.

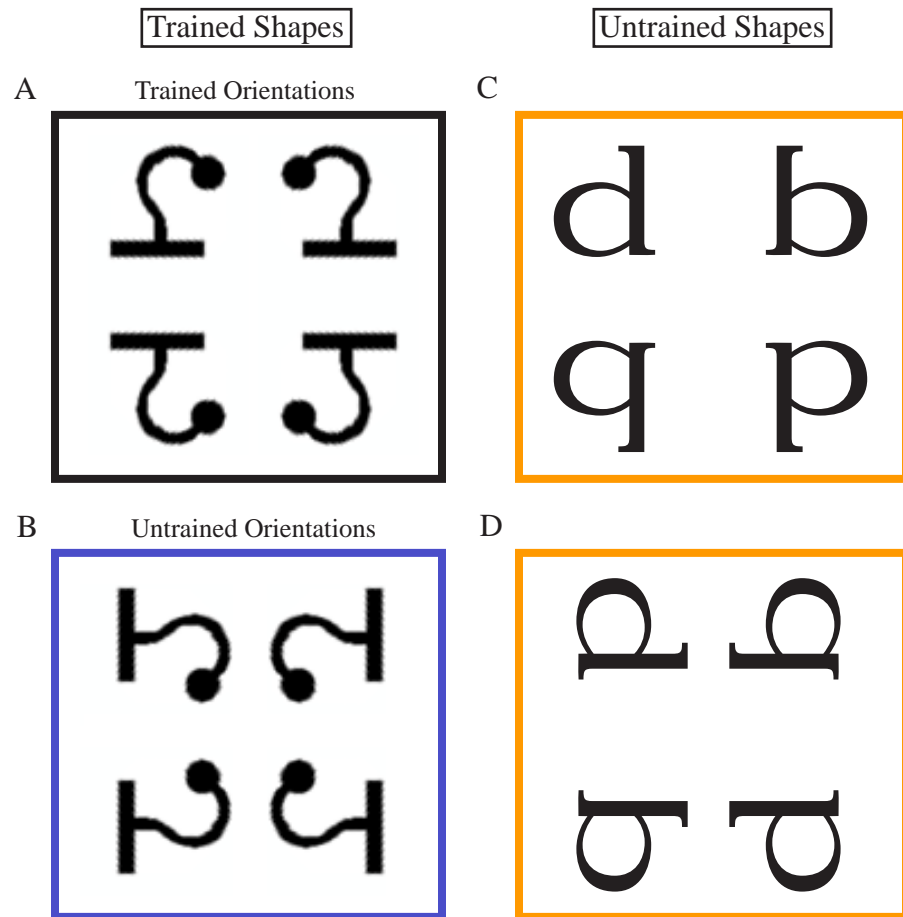


Figure 15. Examples of orientations of trained and untrained shapes presented in the fixation task. A. Four orientations seen in discrimination training by monkey Ph are shown in the black square. Images were formed by rotating the base shape 180° laterally in depth, vertically in depth and in plane. B. Untrained orientations of a trained shape are shown in the blue square. These images were formed by rotating the base shape 90° in plane and then rotating this image 180° laterally in depth, vertically in depth and 180° in-plane. C. Four orientations of an untrained base shape. D. Four orientations of the 90° rotated untrained base shape.

### 3.2.4 Data Collection

At the beginning of each recording session, a neuron in IT was isolated and the twenty shapes were presented foveally to the fixating monkey. One trained and one untrained shape were identified that elicited strong responses from the isolated neuron, as assessed by evaluating the audio monitor output and online histograms. These two shapes were used in the following data collection tasks.

DMS Task: This task was identical to that used in training. Only trained shapes at trained orientations were presented in this task. Neuronal data were collected until the monkey had correctly performed twelve trials per condition.

Fixation Task: This data collection task was used to compare neural responses to trained and untrained shapes. Events in the fixation task during a representative trial are presented in Figure 14. The trial began with the monkey fixating a  $0.6^\circ$  blue fixation spot for 600 ms. An image was then superimposed over the fixation spot for 600 ms. Following offset of the shape, the monkey was required to fixate for an additional 600 ms. The monkey was rewarded with a drop of juice for maintaining fixation within a  $2^\circ \times 2^\circ$  invisible window throughout the entire trial. The images were eight orientation variants of the trained shape (four of them the training orientations) and eight orientation variants of the untrained shape. An example set of images for monkey Ph is shown in Figure 15. The eight orientations of the trained shape formed two tetrads. The first was the training set (Figure 15A). The second tetrad consisted of untrained orientations, rotated  $90^\circ$  in-plane from the trained orientations (Figure 15B). Eight orientations of the untrained shape were formed in the same manner (Figure 15C-D). Images in each set of

eight, formed four lateral mirror image pairs, four vertical mirror image pairs and four 180° in-plane rotations.

In monkey Ph, all sixteen orientations were randomly interleaved. In monkey Op, trained and untrained shapes were run in separate blocks and within each block, the eight orientations were randomly interleaved. The order in which the two blocks were run was varied from session to session. Neuronal data were collected until the monkeys had successfully completed twelve trials of each of the sixteen conditions.

### *3.2.5 Statistical Analysis*

In the DMS task, trials were divided into four epochs for the analysis of neural activity. The “baseline” period consisted of the 400 ms immediately preceding sample onset. The “sample” period began 50 ms after stimulus onset, with a duration of 450 ms. The first 50 ms were excluded from analysis to account for the visual response latency of IT neurons. The time window was designed to include the period of strongest visual responsiveness. The results of the experiment were not sensitive to the exact time window used. The “delay period” began 300 ms after the stimulus offset and extended for 200 ms. The final portion of the delay period was selected for analysis to avoid any visual response to the sample offset. To examine neuronal activity within the “search period”, activity was aligned on saccade onset which was calculated as the time, prior to the monkey fixating the first probe, when eye velocity exceeded 50 degrees per second.

In the fixation task, trials were divided into two epochs for the analysis of neural activity. The “baseline” period consisted of the 400 ms immediately preceding stimulus onset. The “stimulus” period began 50 ms after stimulus onset and extended for 450 ms.



All analyses of neuronal data collected during the DMS and fixation tasks considered data from correct trials only.

Neurons were initially assessed for visual responsiveness in the following manner. Response rates during the baseline and sample periods of the DMS task or the baseline and stimulus periods of the fixation task were compared, for a given shape at a given orientation, using a matched pairs *t*-test, evaluated at  $p < 0.01$ . If this test was significant for at least one of the images in the set, then the neuron was included in the database for subsequent analysis.

For analyses directly comparing response rates between neurons, the square root transform of the average response rates was used. The distribution of response rates tends to follow a Poisson distribution with variances proportional to the mean. The square root transform helps to normalize the firing rate distributions (Zar, 1999). The transform was accomplished with the formula  $X' = (X + 0.5)^{0.5}$ , where  $X$  is the raw firing rate and  $X'$  is the square root transformation (Baker et al., 2002; Vogels and Orban, 1994a).

The index of selectivity (*SI*) used in the trained vs. untrained shape, trained vs. untrained orientation and DMS vs. fixation task comparisons was calculated as  $SI = (b - w) / (b + w)$ , where  $b$  equals the average response rate elicited by the best orientation in a tetrad and  $w$  equals the average response rate elicited by the worst orientation in the tetrad. The “best” orientation was the orientation that elicited the highest response from the neuron and the “worst” orientation was the orientation that elicited the lowest response.

To compare selectivity for lateral and vertical mirror images, *t*-tests were carried out to determine how many image pairs in each category elicited significantly different responses. For each pair of mirror images, for each neuron, the neuron was considered to discriminate between the images if the elicited firing rates were significantly different ( $p < 0.05$ ).

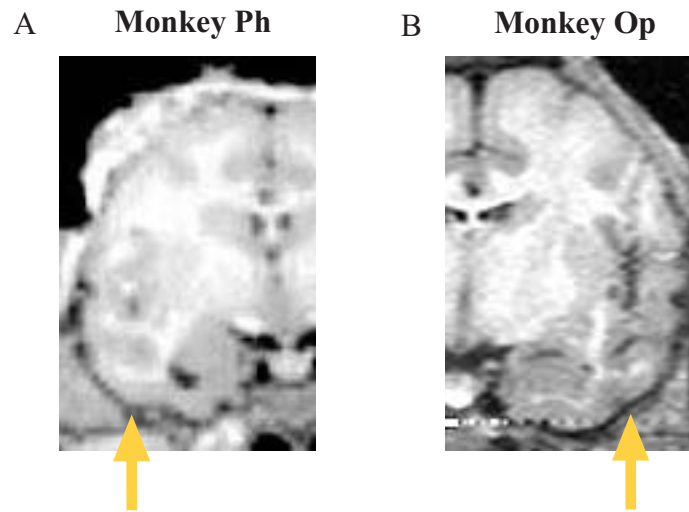
### **3.3 Results**

#### *3.3.1 Overview*

In the following sections, we will describe the results of experiments aimed at characterizing the effects of orientation discrimination training on the responses of IT neurons. In particular, we will begin by describing behavioral data collected during the performance of the DMS task both during training and during neurophysiological testing (section 3.3.3). We will then describe the selective properties of neural data collected during the performance of this task (section 3.3.4). Next we will show that this neuronal selectivity correlates with the behavior of the monkey (section 3.3.5). To determine whether training affected the response properties of IT neurons, we compared neuronal responses to trained and untrained images. Specifically, we will describe comparisons between responses to trained and untrained shapes (section 3.3.6), and between responses to trained and untrained orientations of trained shapes (section 3.3.7). Finally we will conclude by examining the effects of task context on responses to trained orientations. In particular, we will describe comparisons between responses to trained orientations presented in the DMS and fixation tasks (section 3.3.8).

### 3.3.2 Location of Recording Sites

Recording was carried out in anterior IT in the left hemisphere of monkey Ph and the right hemisphere of monkey Op (Figure 16). Recording sites were at frontal levels in the range of anterior 13-16 mm (monkey Ph) and anterior 18-22 mm (monkey Op) as defined with respect to the interaural plane. In both monkeys, recording sites were confined to cortex lateral to the anterior medial temporal sulcus. With respect to depth, recording sites in monkey Ph were localized to the ventral bank of the superior temporal sulcus and the ventral aspect of the inferotemporal gyrus, while recording sites in monkey Op were located on the ventral and lateral aspects of the inferotemporal gyrus.



**Figure 16. Coronal magnetic resonance images.** Images show recording locations in the left hemisphere of monkey Ph (A) and in the right hemisphere of monkey Op (B). Arrows indicate the centers of the recording zones in the two monkeys. In the image from monkey Op, guide tube tracks are visible in the overlying tissue. Recording locations were located at frontal levels in the range of anterior 13-16 mm (monkey Ph), anterior 18-22 mm (monkey Op) and confined to cortex lateral to the anterior medial sulcus (area TE).

### 3.3.3 DMS Task: Behavior

#### Behavior During Training Period

Behavioral data from training sessions prior to electrophysiological recording are shown for each monkey by shape in Figure 17. Both monkeys were trained until they performed above 65% on all shapes (chance = 25%). Monkey Ph required a median of 1163 trials (range 447-2105) to reach criterion and monkey Op required a median of 522 trials (range 159-1332). Monkey Ph was trained for approximately four months and monkey Op for approximately five months.





















#### Behavior During Data Collection Period

During the period of electrophysiological recording, behavioral data were collected from 61 sessions in monkey Ph and 51 sessions in monkey Op.

*Accuracy.* Overall, both monkeys performed well, averaging 78.9% correct and 80.6% correct, for monkeys Ph and Op, respectively. The performance of each monkey is broken down by shape in Table 3. Both monkeys performed well above chance for all shapes (Table 3, column C) and percent correct data did not vary significantly between the two monkeys (Mann Whitney U,  $U = 46.00$ ,  $p > 0.7$ ).

*Confusion Patterns.* We examined the few errors that the monkeys did make to determine whether they tended to exhibit a stereotyped error pattern. We first set out to determine whether there was any systematic pattern of errors across all twelve sample-probe orientation pairs (4 samples x 3 incorrect probes). The results of Chi squared analyses on the frequencies of errors for each of the twenty shapes are shown in Table 3, column G. Error patterns were significantly systematic for all shapes in both monkeys, indicating that the monkeys tended to make certain errors more than others. We then

**Table 3. Behavior of monkeys during electrophysiological recording**

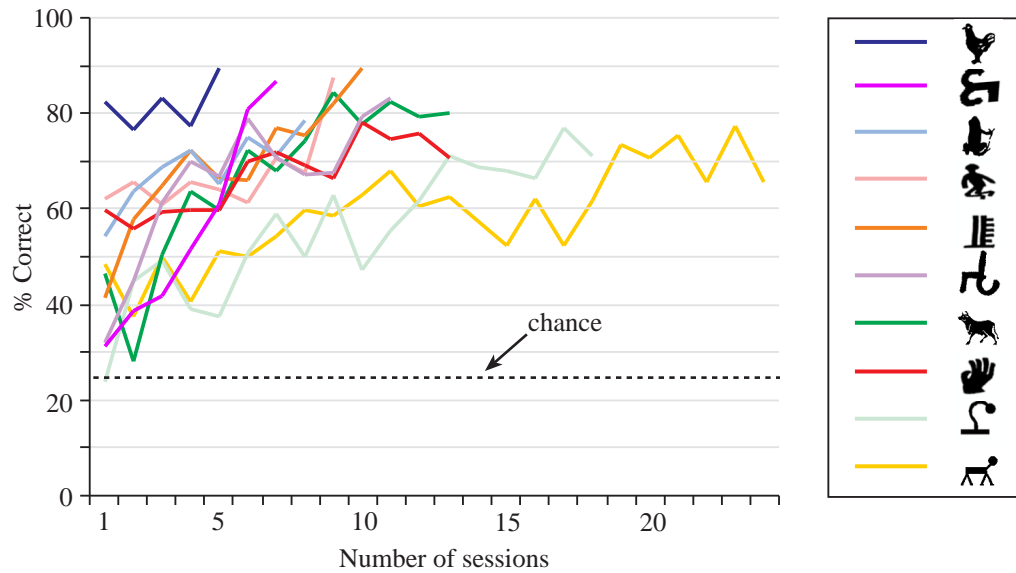
	A	B	C	D	E	F	G	H	I
	Shape	# of sessions	% correct	% lateral errors	% vertical errors	% 180° errors	Sys?*	Lateral Vertical 180°†	Comm?‡
Monkey Ph		2	76.4	45.5	34.8	19.6	<b>0.0005</b> (33.20)	<b>0.2019</b> (3.20)	<b>0.0078</b> (17.45)
		5	62.5	71.1	12.7	16.2	<b>&lt;0.0001</b> (178.52)	<b>&lt;0.0001</b> (102.43)	<b>0.0004</b> (24.79)
		13	80.7	67.4	20.6	12.0	<b>&lt;0.0001</b> (129.08)	<b>&lt;0.0001</b> (80.46)	<b>0.0002</b> (25.80)
		3	77.0	52.6	28.7	18.7	<b>0.0144</b> (23.64)	<b>0.0187</b> (7.95)	<b>0.0542</b> (12.37)
		6	90.7	60.9	22.5	16.6	<b>0.0048</b> (26.89)	<b>0.0075</b> (9.78)	<b>0.5122</b> (5.25)
		5	88.8	23.1	27.5	49.5	<b>0.0003</b> (34.27)	<b>0.5292</b> (1.27)	<b>0.0178</b> (15.33)
		4	71.7	91.4	3.8	4.8	<b>&lt;0.0001</b> (136.61)	<b>&lt;0.0001</b> (116.41)	<b>0.1306</b> (9.86)
		10	77.1	42.7	28.3	29.0	<b>0.0114</b> (24.34)	<b>0.0172</b> (8.13)	<b>0.1914</b> (8.70)
		9	84.4	77.8	16.4	5.8	<b>&lt;0.0001</b> (151.67)	<b>&lt;0.0001</b> (67.57)	<b>&lt;0.0001</b> (34.05)
		4	79.4	36.1	40.8	23.1	<b>&lt;0.0001</b> (52.77)	<b>0.1738</b> (3.50)	<b>0.4701</b> (5.59)
Monkey Op		6	79.0	6.4	83.0	10.6	<b>&lt;0.0001</b> (113.92)	<b>&lt;0.0001</b> (105.69)	<b>0.2805</b> (7.46)
		2	81.6	5.6	94.4	0	<b>&lt;0.0001</b> (78.20)	<b>&lt;0.0001</b> (44.24)	<b>0.1638</b> (9.18)
		12	91.1	78.4	8.2	13.4	<b>&lt;0.0001</b> (130.51)	<b>&lt;0.0001</b> (60.40)	<b>0.3921</b> (6.28)
		5	79.1	96.9	3.1	0	<b>&lt;0.0001</b> (220.18)	<b>&lt;0.0001</b> (120.36)	<b>0.1346</b> (9.77)
		2	86.8	0	100	0	<b>&lt;0.0001</b> (40.20)	<b>&lt;0.0001</b> (30.00)	<b>0.7722</b> (3.29)
		13	84.9	14.5	74.6	10.9	<b>&lt;0.0001</b> (293.46)	<b>&lt;0.0001</b> (86.76)	<b>0.2293</b> (8.12)
		1	75.8	0	75.0	25.0	<b>0.0001</b> (36.50)	<b>0.0009</b> (14.00)	<b>0.7306</b> (3.60)
		3	74.4	0	100	0	<b>&lt;0.0001</b> (199.00)	<b>&lt;0.0001</b> (160.00)	<b>0.6289</b> (4.35)
		3	74.7	6.9	80.2	13.0	<b>&lt;0.0001</b> (50.20)	<b>&lt;0.0001</b> (44.40)	<b>0.3753</b> (6.44)
		4	78.1	0	100	0	<b>&lt;0.0001</b> (129.14)	<b>&lt;0.0001</b> (112.00)	<b>0.9746</b> (1.24)

\* Chi square tests done on twelve sample-choice pairs to determine whether errors were systematic. Bold numbers indicate *p* values with Chi squared values in parentheses.

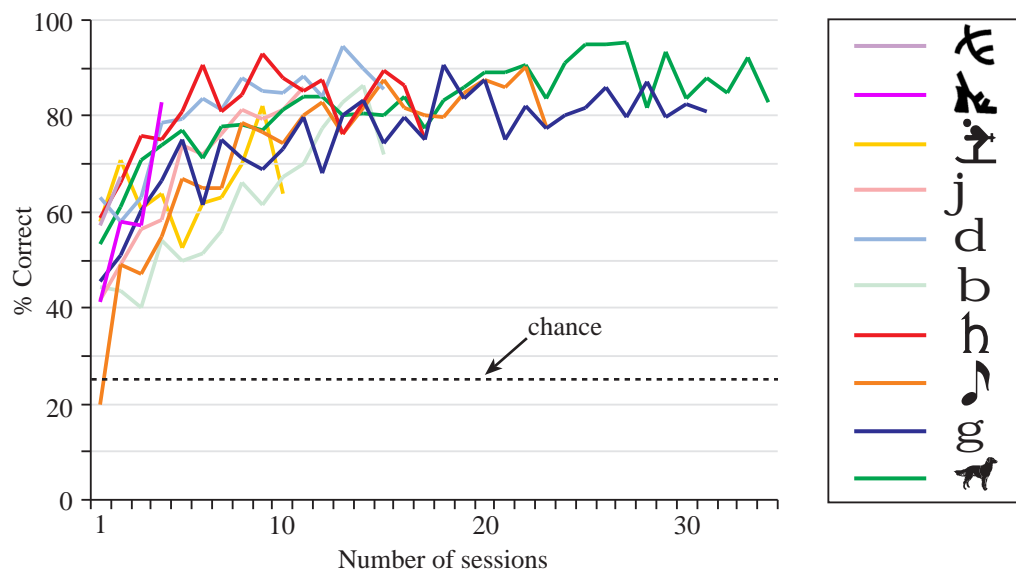
† Systematicity of error patterns in columns D-F

‡ To determine whether errors were commutative, two sets of six sample-choice pairs were compared using Mc Nemar's Change Test.

### A. Monkey Ph



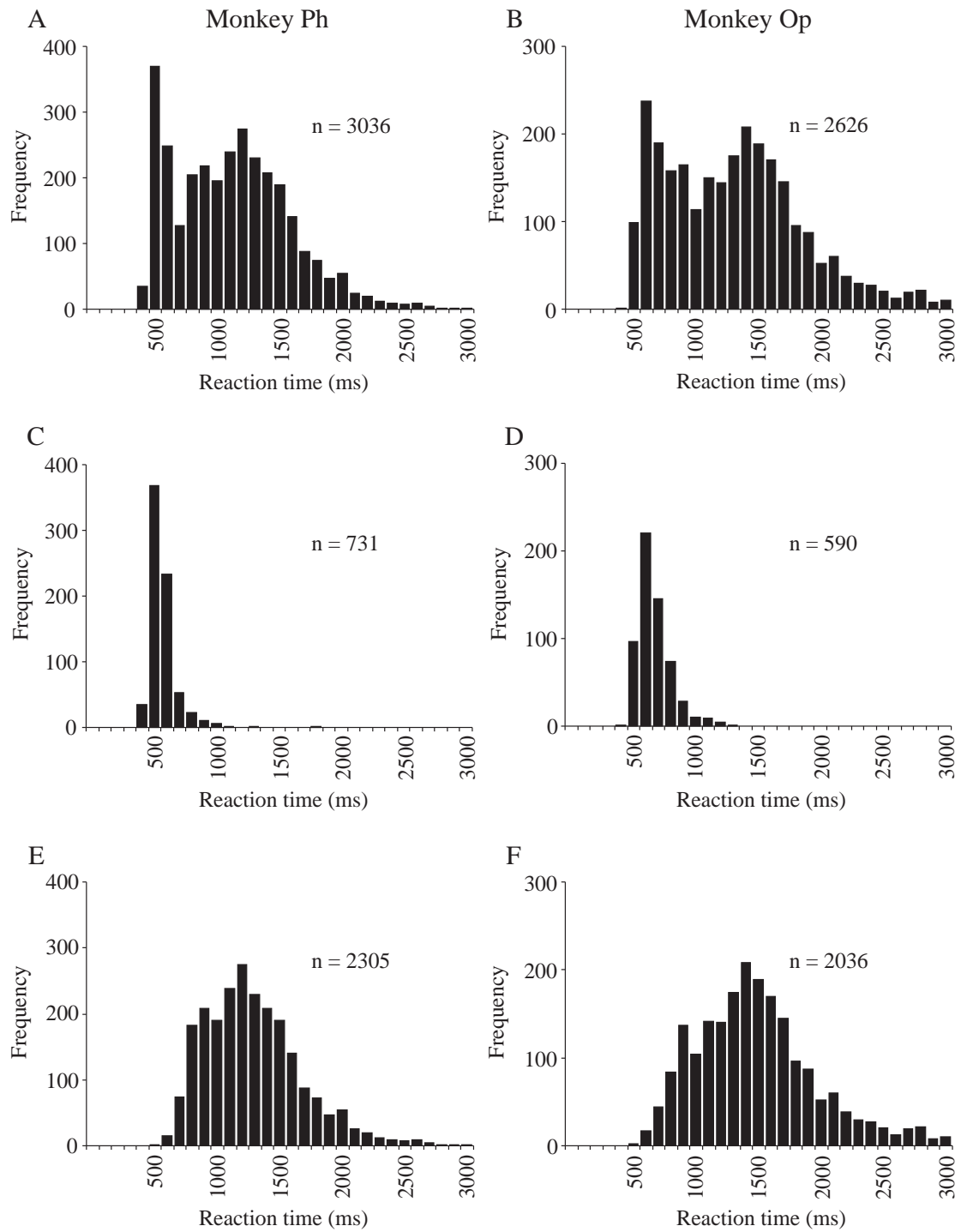
### B. Monkey Op



**Figure 17. Learning curves for each of ten shapes seen in the DMS task during training.** Data are shown separately for monkey Ph (A) and monkey Op (B). Monkeys were trained until they performed at least 65% correct on all ten shapes (chance = 25%). In monkey Ph, this required a median of 1163 trials per shape, and in monkey Op this required a median of 521 trials per shape.

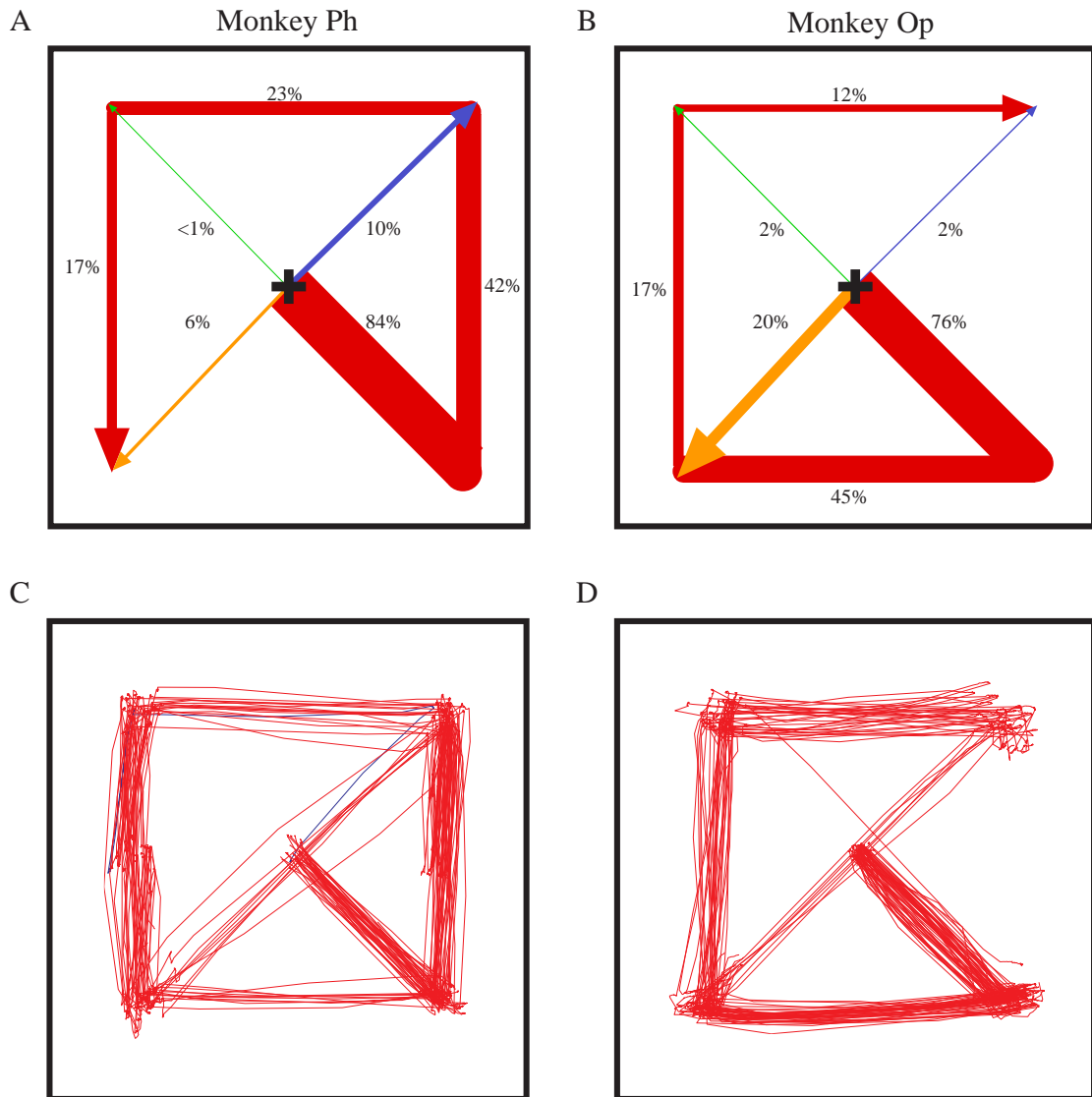
asked whether the monkeys tended to confuse certain kinds of pairs more than others. We counted instances in which the sample's lateral or vertical mirror image or its 180° rotation was selected erroneously. The frequencies of these three error types are shown in Table 3, column D-F. The pattern of errors was clearly systematic with respect to the type of the pair. A Chi squared analysis (column H) revealed this to be the case for most shapes considered. It is clear from these data that the frequencies of lateral and vertical mirror image errors differed between the two monkeys. Specifically, monkey Ph made primarily lateral mirror image errors for eight out of ten images. In contrast, monkey Op made primarily vertical mirror image errors for eight out of the ten trained images. Finally, we explored whether errors were commutative, that is, if orientation A is given as the sample and the monkey chose orientation B, did the monkey tend to choose orientation A when given orientation B as the sample. Frequencies of errors were analyzed using McNemar's change test. Results of this analysis are given in column I of Table 3 for each shape. For both monkeys, errors tended to be commutative for most shapes.

*Reaction Time.* The distributions of reaction times for the two monkeys are shown in Figure 18A-B. Average reaction times calculated from probe onset to lever press were 1056 ms and 1230 ms for monkey Ph and Op, respectively. For both monkeys, the distributions of reaction times appeared to be bimodal. One possible reason for this is that the initial peak in reaction times corresponds to trials in which the monkey fixated the match orientation first and responded immediately, while the second peak in reaction times corresponds to trials in which the monkey made multiple fixations before responding. To determine if this was the case, trials were divided into two groups:



**Figure 18. Reaction times in the DMS task.** Reaction times were calculated from probe onset to lever press for 61 sessions run in monkey Ph (A,C,E) and 51 sessions run in monkey Op (B,D,F). A-B. Reaction times from all correct trials. C-D. Reaction times from correct trials in which monkeys fixated only one probe before responding. E-F. Reaction times from correct trials in which monkeys fixated more than one probe before responding.





**Figure 19. Scan path of monkeys during search period of DMS task.** A-B. Summary of eye data from 61 sessions in monkey Ph (A) and from 51 sessions in monkey Op (B). Red, orange, green and blue arrows radiating from the center represent saccades from the fixation spot to probes at bottom right, bottom left, upper left and upper right locations, respectively. Width of arrows indicates relative proportion of correct trials in which the initial saccade was toward that specific location. Most common second, third and fourth saccades are also shown, with line width representing proportion of trials in which saccades had followed that scan path up to and including that segment. C-D. Eye traces from a representative session from monkey Ph (C) and monkey Op (D). Red, orange, green and blue traces represent eye traces during trials in which the monkey made his first saccade following probe onset to the bottom right, bottom left, upper left and upper right locations.

correct trials in which the monkey fixated only one probe and then responded (Figure 18C-D), and correct trials in which the monkey fixated more than one probe before responding (Figure 18E-F). The two distributions of reaction times for each monkey clearly correspond to the two peaks in the bimodal distributions in (A) and (B), indicating that the initial peak in the combined distribution represents responses following a single fixation.

*Scanning Patterns.* We next analyzed the scan path of the monkeys during the search period to determine whether saccades followed a set spatial pattern or were instead driven by the identity of the probes (as would occur if the monkey detected the match orientation with peripheral vision and then looked directly at it). A summary of the data is shown in Figure 19 for monkeys Ph (A) and Op (B). The relative width of arrows radiating from the center represents the relative numbers of initial saccades across all correct trials from all sessions directed at that location. For both monkeys, the majority of initial saccades were toward the lower right probe (red arrow). If the match was not found at this location, each monkey engaged in a stereotyped search pattern (counterclockwise for monkey Ph and clockwise for monkey Op). The successive segments of the red trajectory in Figure 19A-B have widths corresponding to the percentage of trials on which the monkey traversed this segment of the trajectory. The widths grow less because search generally stopped when the monkey encountered the match probe. Examples of scanning patterns from representative recording sessions for the two monkeys are presented in Figure 19C-D. As in Figure 19A-B, scan paths initially directed to the lower right probe are shown in red. For both monkeys, in almost all trials, monkeys made their initial saccade to the lower right probe and continued in

their stereotyped scan path if the match was not at this location. In only one trial, in monkey Ph, was the initial saccade directed elsewhere (blue line, Figure 19C).

### *3.3.4 DMS Task: Neural Activity*

#### Sample Period Activity

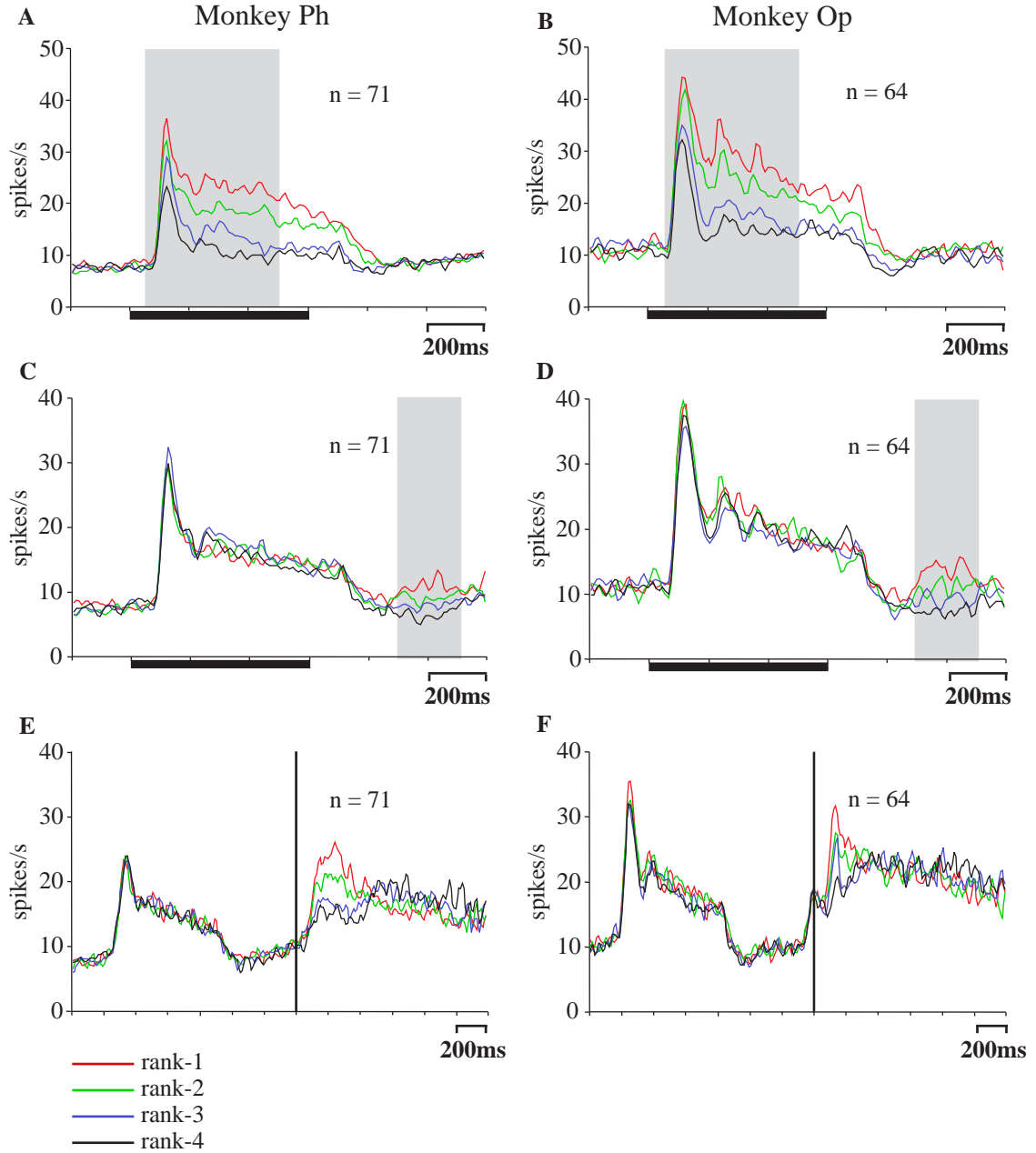
Of the 185 neurons (monkey Ph:  $n = 100$ ; monkey Op:  $n = 85$ ) studied during performance of the DMS task, 135 (monkey Ph:  $n = 71$ ; monkey Op:  $n = 64$ ) responded to at least one of the four orientations in the shape set selected for testing. In order to determine the time course of responses to the four trained orientations, population histograms were created for these neurons. Specifically, for each neuron, the four orientations were ranked according to the strength of the response evoked during the sample period. Then, responses to each rank were averaged across all neurons. The resulting population histograms are presented in Figure 20 for monkey Ph (A) and monkey Op (B). In these histograms, neuronal activity is selective during the epoch in which firing rate determined the identification of rank-1, 2, 3 and 4 orientations (gray bars in Figure 20A-B). Selective activity persisted until around 200 ms after shape offset and then subsided. An analysis of variance (ANOVA) performed on the data from these neurons during the same measurement epoch revealed that 97 of the 135 visual neurons (monkey Ph:  $n = 52$ ; monkey Op:  $n = 45$ ) exhibited significant selectivity during the sample period.

### Delay Period Activity

It is evident in the population histograms shown in Figure 20A-B that selective activity did not persist into the later delay period. This suggests that IT neurons did not carry the trace on which the monkey's subsequent response was based. To investigate this issue further, we constructed population histograms for shapes ranked 1, 2, 3 and 4 on the basis of neuronal activity later in the delay period (300-500 ms following image offset). In these histograms, neurons exhibited selectivity during the measurement epoch (gray bars in Figure 20C-D), as they must be by definition. Selectivity during the delay period did not, however, appear to be correlated with selectivity during the sample period. That is, responses during the sample period do not assume the same rank order as responses during the delay period. Of the 135 visual neurons studied, only 8% (11/135) had significant selectivity during the delay period (monkey Ph: 1/71; monkey Op: 10/64). This fraction of neurons is not significantly more than the fraction expected by chance from type 1 errors ( $\chi^2 = 1.97$ ,  $p > 0.1$ ). We conclude that few neurons exhibited selectivity during the delay period of the DMS task. Thus, IT neurons were selectively active in conjunction with visual processing of the trained orientations but not in conjunction with holding them in working memory.

### Search Period Activity

We next asked whether neural activity during the search period correlated with the particular orientation the monkey was fixating. To address this question, we created population histograms of neural activity from visual neurons recorded from during the DMS task, with activity aligned on saccade onset. Individual trials were grouped



**Figure 20. Population histograms of neural activity collected during performance of the DMS task.**

Data are shown separately for monkey Ph (A,C,E) and monkey Op (B,D,F). A-B. Population activity elicited by four trained orientations during the sample period. For each neuron, the four orientations were ranked according to response strength during the sample period (gray region). Activity as a function of time was averaged across all neurons for each rank. C-D. Average population responses elicited by four trained orientations during the delay period. For each neuron, the four orientations were ranked according to response strength during the last 200 ms of the delay period (gray region). Activity as a function of time was averaged across all neurons for each rank. E-F. Average population responses elicited by four trained orientations during the search period. For each neuron, trials were sorted by the orientation the monkey first fixated following probe onset. These four groups of trials were ranked according to the response strength evoked by that orientation during the sample period. Activity as a function of time was averaged across all neurons for each rank. Activity was aligned on saccade onset (vertical line). Black bar in A-D represents time of sample presentation. Red line: rank 1; green line: rank 2; blue line: rank 3; black line: rank 4. Curves were smoothed according to the formula  $Y_n' = 0.25Y_{n-1} + 0.5Y_n + 0.25Y_{n+1}$ , where  $Y_n$  is instantaneous firing rate.

according to the first orientation the monkey fixated following probe onset. The four groups were ranked according to response rates evoked by the orientation during the sample period. Responses were then averaged across all neurons for each of the four ranks with a temporal resolution of 10 ms. The resulting population histograms are shown in Figure 20 for monkey Ph (E) and monkey Op (F). Trials in which, following probe onset, the monkey fixated the orientations ranked first, second, third and fourth are represented by the red, green, blue and black lines respectively. It is evident that following initiation of the first saccade (alignment event indicated by the vertical lines in Figure 20E-F) neuronal activity varied by shape rank even though the ranking of shapes had been determined by firing rate during the sample period (Figure 20A-B). In other words, IT neurons exhibited the same pattern of visual selectivity during fixation of the probe as they had during fixation of the sample. Two points are worth note. First, neuronal activity was not selective for the identity of the probe immediately prior to initiation of the saccade. This may reflect the fact that probe discriminability was reduced in peripheral vision. Second, selective activity did not persist long. This is related to the fact that fixation was generally brief. On average, monkeys fixated the first probe for 144 ms (monkey Ph) and 161 ms (monkey Op). Late in the trial, the firing rate pattern actually reverses, with neurons firing at the lowest rate when the probe fixated first had been the rank-1 shape (red lines in Figure 20E-F). This reflects the fact that probes visited during later stages of search were those other than the rank-1 probe. From this analysis, we conclude that neural activity during the initial portion of the search period reflects the shape orientation that the monkey is fixating.

### 3.3.5 DMS Task: Correlation Between Behavioral Errors and Neuronal Activity

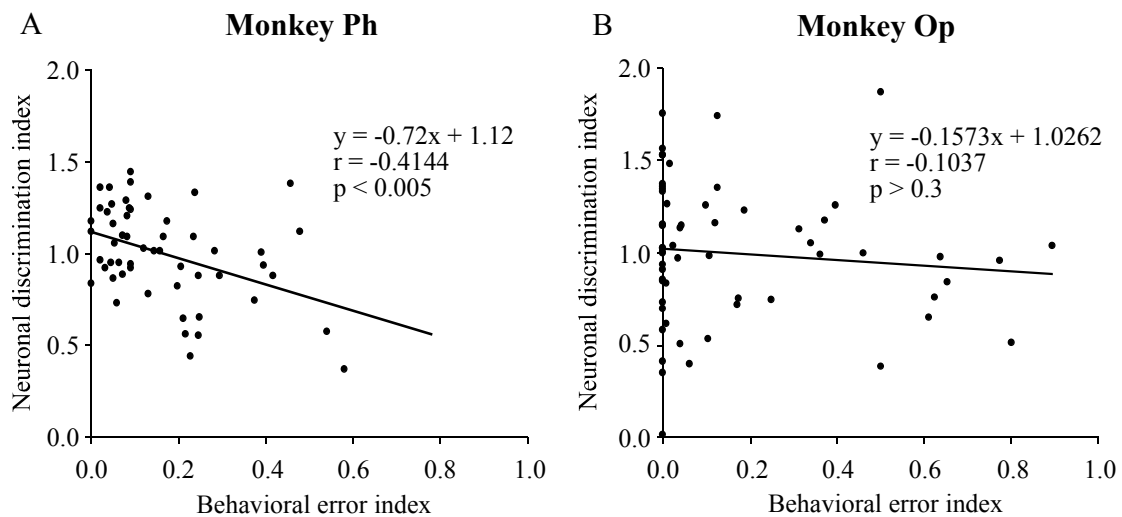
We next investigated whether the tendency of neurons in IT to confuse shapes was correlated with the tendency of the monkey in the DMS task to confuse them. To test this, the selectivity of IT neurons for pairs of orientations presented during the sample period of the DMS task was compared to the ability of the monkey to discriminate between the same pair of orientations. This was done by calculating a neuronal selectivity index and a behavioral error index for each of the six orientation pairs for each shape seen in the DMS task (two lateral, two vertical and two 180° rotation pairs as in Figure 13B). The neuronal selectivity index was calculated as the absolute difference in average firing rates elicited by the two orientations presented during the sample period, divided by the average difference in firing rates across all six pairs. This index was calculated for each of the six pairs for each neuron and then the values for each pair were averaged across all neurons tested with a given shape, resulting in 60 values of neuronal orientation selectivity for each monkey (six pairs x ten trained shapes). The behavioral error index was the number of errors the monkey made for a given pair divided by the total number of errors made for all six pairs. This index was calculated for each of the six pairs for each of the trained shapes<sup>\*†</sup>.

If the orientation selectivity of IT neurons were related to behavioral discrimination, we would expect the relation to be manifested in a negative correlation between the neuronal selectivity index and the behavioral error index. That is, the more selective IT neurons were for a given pair of orientations (high selectivity index), the

<sup>\*</sup> There were too few errors to calculate behavioral error indices by neuron, so errors were counted across all recording sessions for each pair of a given shape.

<sup>†</sup> Only nine trained shapes were included for monkey Ph because one of the ten shapes failed to elicit significant visual responses from the neurons studied.

better the monkey would be at discriminating between the two orientations (low behavioral error index). Figure 21A-B shows the behavioral error indices plotted against the neuronal orientation selectivity indices for the 71 and 64 visual neurons recorded in the two monkeys during the DMS task. Regression lines are superimposed over each set of points. In monkey Ph, there was a clear negative correlation between neuronal selectivity and behavioral errors (Pearson product-moment correlation:  $r = -0.4144$ ,  $p < 0.005$ ). In monkey Op, the trend was negative, but did not attain significance ( $r = -0.1037$ ,  $p > 0.3$ ). The degrees of correlation for the two monkeys were not significantly different ( $t = 1.749$ ,  $p > 0.08$ ). This justified combining the data for analysis, and doing so revealed a significant correlation ( $r = -0.1863$ ,  $p < 0.05$ ).



**Figure 21. Correlation between neuronal orientation selectivity and behavioral discrimination of orientation in the DMS task.** Data are shown separately for monkey Ph (A) and monkey Op (B). Each point represents data from a single pair of trained orientations of a given shape. The four trained orientations form six possible pairs (see Figure 13B). Values on abscissa are measures of the behavioral error index calculated as the number of errors for a given pair divided by the total number of errors for all six pairs of a given shape. Errors were counted across all sessions. Values on the ordinate are measures of the neuronal discrimination index calculated as the difference in response rates evoked by the two members of each pair of orientations divided by the average difference in response rates for all six pairs of a given shape. Index values were calculated for each of the six pairs for each neuron and then values for each pair were averaged across all neurons tested with a given shape. Regression lines are superimposed over the data. There was a significant negative correlation in monkey Ph ( $p < 0.005$ ) and a trend toward a negative correlation in monkey Op ( $p > 0.3$ ). The correlations for the two monkeys were not significantly different ( $p > 0.8$ ) and analysis of the combined data revealed a significant correlation ( $p < 0.05$ ).

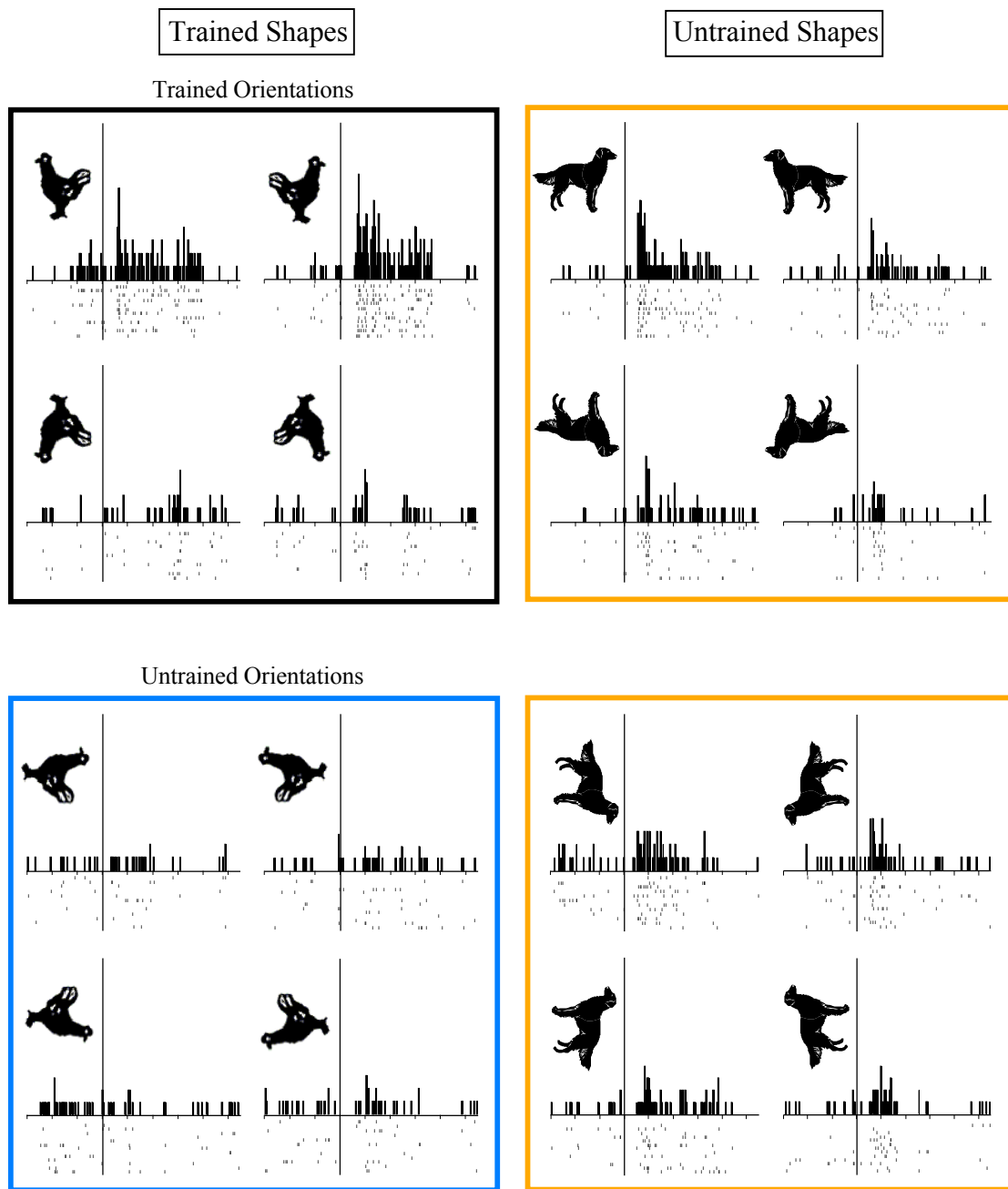


### 3.3.6 Fixation Task: Trained vs. Untrained Shapes

In order to determine the effects of orientation training on IT neurons, we compared responses of neurons to trained shapes and untrained shapes in the context of the fixation task. An example of activity recorded from one neuron during performance of the fixation task is shown in Figure 22. This neuron exhibited stronger selectivity among the four orientations in the trained tetrad (black box) than among orientations of untrained shape (orange boxes). Furthermore, the strongest response was to a trained shape. To test whether this was true for the population of IT neurons studied, we carried out analyses comparing responses evoked by trained orientations to responses evoked by untrained shapes. Specifically, we compared responses of neurons to trained and untrained shapes presented in the fixation task in terms of response rate, degree of overall selectivity and the degree to which members of lateral and vertical mirror image pairs were discriminated. Data were collected from 180 IT neurons (monkey Ph:  $n = 104$ ; monkey Op:  $n = 76$ ). Of these neurons, 175 were determined to be visually responsive (monkey Ph:  $n = 99$ ; monkey Op:  $n = 76$ ). Only these visually responsive neurons were considered in the following analyses.

#### Effects of Training on Firing Rate

We first asked whether training influenced the overall response strength of the neurons. The firing rate elicited by the preferred shape in the trained tetrad was compared to the firing rate elicited by the preferred shapes in each of the untrained



**Figure 22. Responses of a single neuron to orientations of trained and untrained shapes presented foveally in the fixation task.** This neuron exhibits higher selectivity for trained orientations of trained shapes than for orientations of untrained shapes. Data were aligned on the onset of the 600 ms shape (vertical line traversing histogram and rasters). Vertical calibration bar represents 100 spikes/s. Tick marks on the horizontal axis are in 200 ms increments. The histogram bin widths were 10 ms.

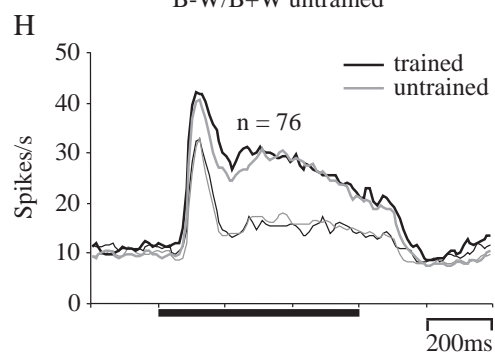
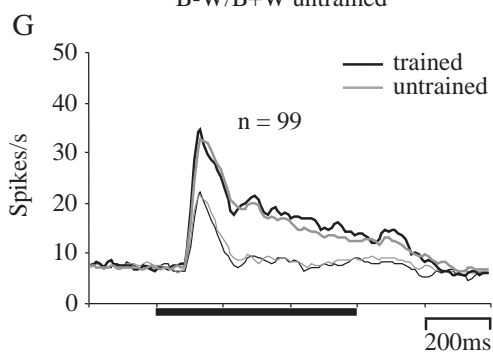
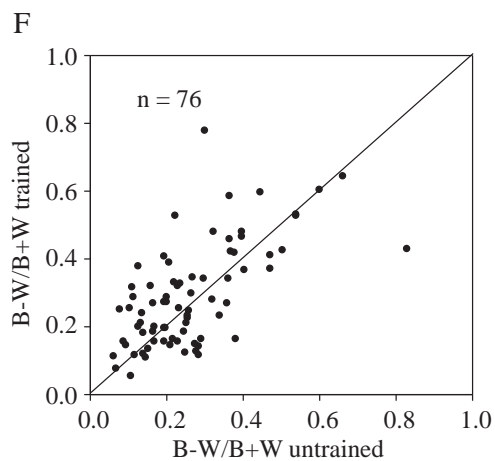
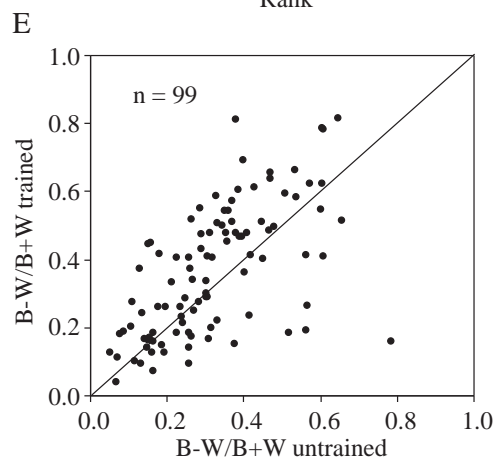
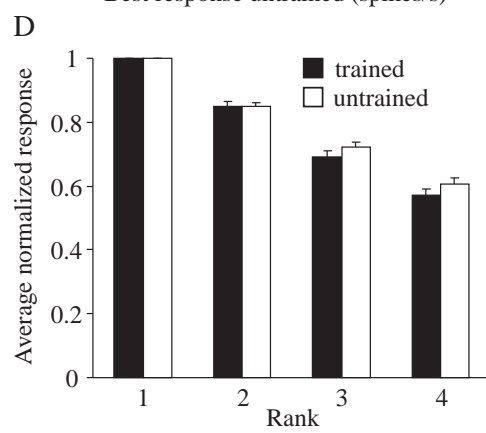
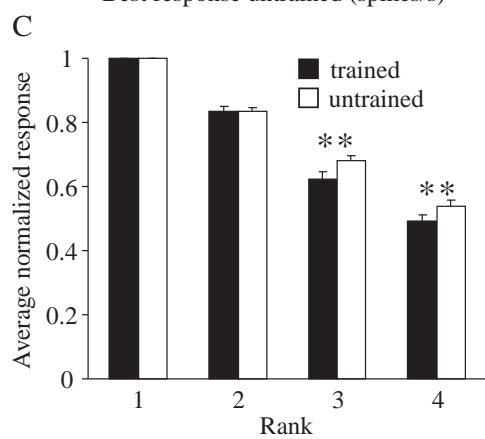
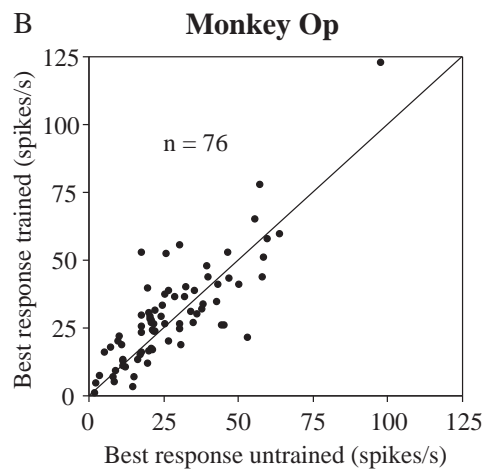
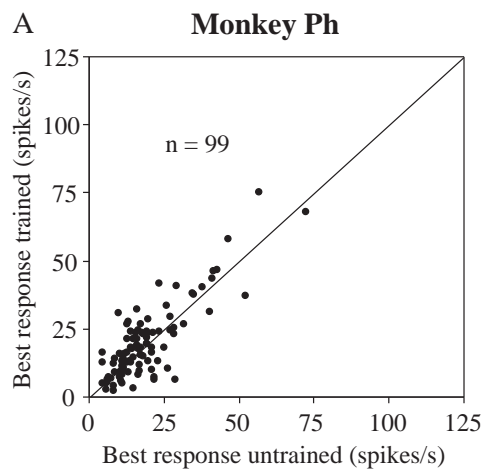
tetrads. For each visually responsive neuron, mean response rates were calculated over the stimulus period (50-500 ms relative to shape onset) for each orientation in the trained tetrad and the two untrained tetrads. The ‘preferred’ shape in each tetrad was then identified as the one eliciting the strongest response. The responses to the preferred shapes in the two untrained tetrads were averaged together. Statistical analyses were then performed on the square root transforms of the data (see methods).

Figure 23 (A-B) shows the mean response rate for the preferred trained image plotted against the mean response rate for the preferred untrained image for each neuron (monkey Ph: mean trained = 19.54 spikes/s, mean untrained = 18.82 spikes/s; monkey

---

**Figure 23. Comparison of neural activity elicited by orientations of trained and untrained shapes presented in the fixation task.** A-B. Responses to the preferred orientations of trained tetrads plotted against the average of the responses to the preferred orientations in the two tetrads of untrained images for monkey Ph (A) and monkey Op (B). Each point represents data from a single neuron. The ‘preferred’ orientation in each tetrad was defined as the one eliciting the strongest response. For untrained shapes, the response strength evoked by the preferred orientations in the two tetrads were averaged together. There was no significant tendency in either monkey for responses elicited by the preferred trained orientation to exceed those elicited by the preferred orientation of untrained shapes (ANOVA,  $p > 0.1$ ). C-D. Comparison of ranked neuronal responses elicited by trained orientations to ranked neuronal responses elicited by orientations of untrained shapes for monkey Ph (C) and monkey Op (D). Bars represent mean normalized response strengths across all neurons for trained orientations (black bars) and for orientations of untrained shapes (white bars). For each neuron, orientations in the trained tetrad were ranked from the most effective to the least effective. The same was done for the two untrained tetrads. Firing rates were then normalized to the firing rate elicited by the best orientation in each tetrad. The normalized firing rates for each rank of the two untrained tetrads were then averaged together. The mean firing rate at each rank of the trained and untrained tetrad was then computed across all sessions. A repeated measures ANOVA revealed a significant effect of training ( $p < 0.006$ ). Error bars represent standard error of the mean (s.e.m.). Asterisks indicate a significant difference between trained and untrained firing rates as determined by a *post hoc* analysis (Tukey HSD,  $**p < 0.005$ ). E-F. Scatter plot of index values representing selectivity within trained and untrained tetrads for monkey Ph (E) and monkey Op (F). Index of selectivity =  $(b - w) / (b + w)$ , where  $b$  and  $w$  are response rates elicited by best and worst orientations in the tetrad. Index values for the two untrained tetrads were averaged together. The distribution of index values across all neurons was significantly shifted toward higher values for trained orientations (ANOVA:  $p < 0.002$ ). G-H. Population histograms of neural activity collected during performance of the fixation task for monkey Ph (G) and monkey Op (H). Black lines represent activity during trials when trained orientations of trained shapes were presented and grey lines represent activity during trials when orientations of untrained shapes were presented. Thick lines represent responses for best orientations and thin lines represent responses for worst orientations. Activity was aligned on shape onset and black bars indicate duration of shape presentation. Tick marks on horizontal axis are in 200 ms increments. Histogram bin widths were 10 ms. Curves were smoothed according to the formula  $Y_n' = 0.25Y_{n-1} + 0.5Y_n + 0.25Y_{n+1}$ , where  $Y_n$  is instantaneous firing rate.

---



Op: mean trained = 29.66 spikes/s, mean untrained = 27.79 spikes/s). To assess whether the distributions differed significantly, we carried out a repeated measures ANOVA, with transformed response rate as the dependent variable, and training status (trained or untrained) and monkey (monkey Ph or monkey Op) as factors. This analysis revealed a significant difference in firing rate between the two monkeys ( $p < 0.00005$ ) but no significant main effect of training status ( $p > 0.1$ ), and no interaction between monkey and training status ( $p > 0.3$ ). To ensure that the order in which trained and untrained blocks were run in monkey Op did not affect the overall response rates\*, data from monkey Op were analyzed separately, using a repeated measures ANOVA with training status as the within subjects factor and block order as the between subjects factor. The results of this analysis confirmed the previous results in that there was no effect of training status ( $p > 0.1$ ). Additionally, there was no effect of block order ( $p > 0.2$ ) and no interaction between training status and block order ( $p > 0.6$ ).

#### Effects of Training on Overall Selectivity:

To determine whether orientation discrimination training increased the orientation selectivity of IT neurons, we examined how quickly responses dropped off from the best to worst orientations for the trained and untrained shapes. Specifically, for each neuron, we ranked the four orientations in the trained tetrad according to the strength of the visual response and did the same for the two untrained tetrads. In order to eliminate any effect of absolute response strength, we then normalized the responses elicited by each of the three low ranked shapes to the response elicited by the best shape in the tetrad.

\* For 26 of the 76 visual neurons, the learned block was run first. For 50 neurons, the unlearned block was run first.

Normalized firing rates for the two untrained tetrads were averaged together at each of the four ranks. The normalized responses for the four ranks averaged across all neurons are shown in Figure 23C-D. For both monkeys, there appears to be a trend toward greater selectivity for trained shapes. That is, firing fell off more sharply from the rank-1 shape to the rank-4 shape in the trained tetrad, indicating that neurons were more sharply tuned for trained shapes.

To assess the significance of this effect, we carried out a repeated measures ANOVA, with normalized response strength as the dependent variable, and training status (trained or untrained) and rank (2-4) as the within subjects factors and monkey (monkey Ph or monkey Op) as the between subjects factor. Rank-1 orientations were excluded from the analysis since normalization eliminated all variance at this rank. Including rank as a factor allowed us to assess whether training effects varied across rank. This analysis revealed an overall significant main effect of training status ( $p < 0.006$ ), a significant main effect of rank ( $p < 0.000001$ ) and an interaction between training status and rank ( $p < 0.0002$ ). *Post-hoc* analyses revealed significant training effects for monkey Ph at rank-3 ( $p < 0.00003$ ) and rank-4 ( $p < 0.0006$ ). While there existed a main effect of monkey ( $p < 0.02$ ), critically, there was no significant interaction between monkey and training status ( $p > 0.5$ ). To ensure that the order in which the separate blocks were run in monkey Op did not influence the outcome of this analysis, data collected from monkey Op were analyzed separately using block order as a between subjects variable. There was no main effect of block order ( $p > 0.6$ ) and no interaction between block order and training status ( $p > 0.2$ ).

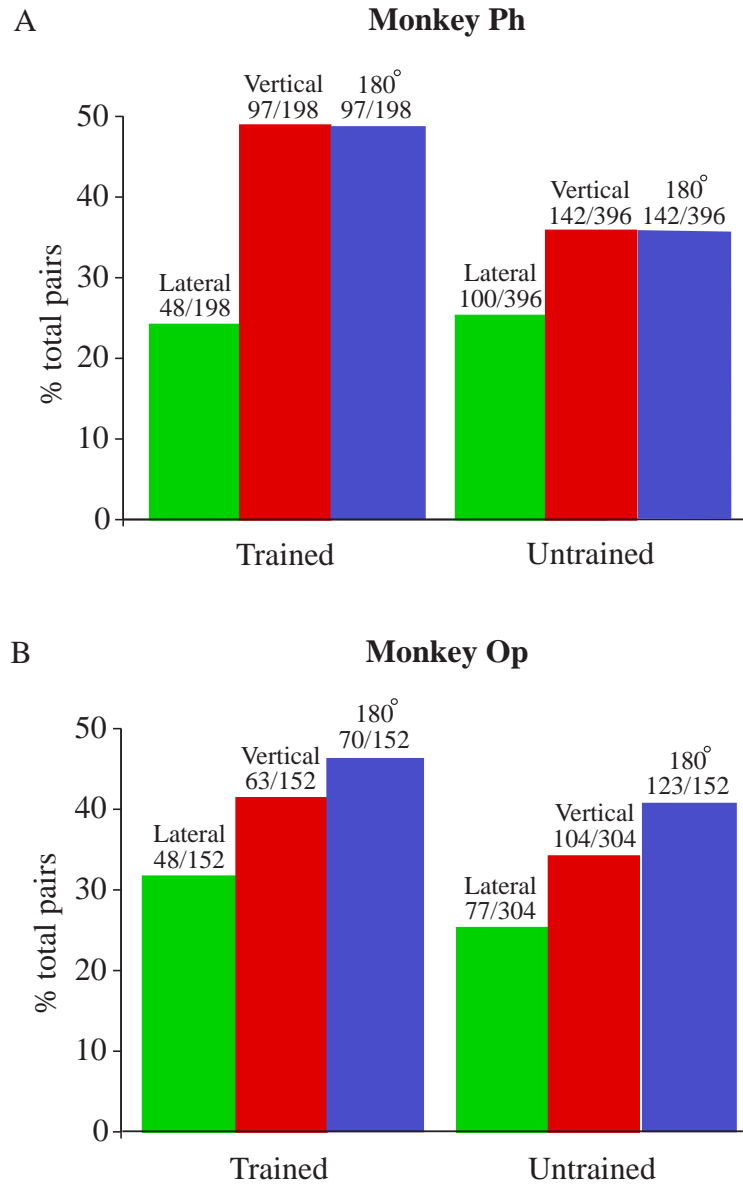
We next asked whether the increase in selectivity for trained shapes was a result of low-ranked trained orientations eliciting significantly weaker responses than low-ranked untrained orientations. To test this, firing rates elicited by the lowest ranked orientations (rank-4) of trained and untrained shapes were compared using a repeated measures ANOVA, with transformed response strength as the dependent variable, training status (trained or untrained) as the within subjects factor and monkey (monkey Ph or monkey Op) as the between subjects factors. This analysis revealed a main effect of monkey ( $p < 0.00002$ ) but no significant interaction between monkey and training status ( $p > 0.4$ ). There was no main effect of training status ( $p > 0.5$ ). The mean response rate for rank-4 trained orientations was less in monkey Ph (mean trained = 9.98 spikes/s, mean untrained = 10.56 spikes/s) but more in monkey Op (mean trained = 17.18 spikes/s, mean untrained = 17.12 spikes/s). These results leave open the question of whether the increase in selectivity for the training set is due to a slight (but not significant) increase in rank-1 responses, a slight (but not significant) decrease in rank-4 responses, or both.

We next sought to determine whether the modest increase in selectivity for trained orientations was the product of large changes in a small subset of IT neurons or small changes in many neurons. To examine this, we calculated selectivity indices for the trained tetrad and for the two untrained tetrads for each neuron (see methods). Indices for the two untrained tetrads were averaged together. Selectivity indices for all visual neurons in monkey Ph and monkey Op are shown in Figure 23E-F. There does not appear to be a subset of neurons exhibiting much higher selectivity for trained than for untrained shapes. Rather, there is a subtle trend affecting the population as a whole. To

assess the significance of this trend, we carried out a repeated measures ANOVA with selectivity index strength as the dependent variable, and training status (trained or untrained) and monkey (monkey Ph or monkey Op) as factors. Overall, there existed a main effect of training status ( $p < 0.002$ ), a main effect of monkey ( $p < 0.004$ ), and no interaction between training status and monkey ( $p > 0.4$ ). Analysis on data from monkey Op alone revealed no significant effect of block order ( $p > 0.7$ ) and no interaction between block order and training status ( $p > 0.4$ ).

To determine the time course of selective responses in IT, we constructed population histograms of the best (rank-1) and worst (rank-4) orientations for trained and untrained shapes. These histograms, constructed separately for the two monkeys, are shown in Figure 23G-H. Black lines indicate responses evoked by trained orientations and grey lines indicate responses evoked by untrained orientations. Thick and thin lines are responses to rank-1 (best) and rank-4 (worst) orientations, respectively. For both monkeys, responses to the best and worst orientations are, for the most part, similar across trained and untrained shapes. There are, however, subtle distinctions. In particular, responses to the best untrained orientations are slightly weaker than those to the best trained orientations. This is evident in the later portion of the response period in monkey Ph (Figure 23G) and in the initial portion of the response period in monkey Op (Figure 23H).





**Figure 24. Frequencies of orientation pairs for both trained and untrained tetrads that elicited significantly different responses from neurons.** Data are shown separately for monkeys Ph (A) and Op (B). Frequencies are shown for lateral mirror image pairs (green bars), vertical mirror image pairs (red bars) and 180° rotations (blue bars). Values along ordinate represent the percent of orientation pairs that elicited significantly different responses out of all possible pairs. Significance was determined by a *t*-test ( $p < 0.05$ ).

### Effects of Training on Neuronal Discrimination Between Lateral and Vertical Mirror Images

We next asked whether training monkeys on the discrimination task differentially affected the selectivity for lateral mirror images and vertical mirror images. One possibility is that training might have reduced the tendency demonstrated in untrained monkeys for neurons to discriminate between vertical mirror images more effectively than between lateral mirror images. To test this, we counted instances in which neuronal activity discriminated between two images in a pair. This was done for each of the two lateral and two vertical mirror image pairs in the trained tetrad and each of the four lateral and four vertical mirror images in the two untrained tetrads. Figure 24A-B shows the frequencies of lateral pairs (green bars) and vertical pairs (red bars) that elicited significantly different responses for trained and untrained shapes in monkeys Ph and Op (frequencies of 180° rotated pairs that elicited significantly different responses are shown in blue for comparison). It is clear in both monkeys that training did not eliminate the tendency for neurons to discriminate between vertical mirror images more often than lateral mirror images. However, in accord with the analyses described above, the overall frequency of pairs discriminated was increased with training. A loglinear analysis revealed that this was indeed the case. There was no main effect of monkey ( $X^2 = 0.00$ ,  $p > 0.9$ ), but the main effect of training status ( $X^2 = 7.53$ ,  $p < 0.007$ ) and image relation (lateral vs. vertical mirror image;  $X^2 = 37.37$ ,  $p < 0.0001$ ) were significant. Critically, there was no interaction between training status and image relation ( $X^2 = 2.20$ ,  $p > 0.1$ )

and no interaction between training status and monkey ( $X^2 = 0.10$ ,  $p > 0.7$ )\*. We conclude that training the monkey to discriminate between lateral and vertical mirror images enhances equally the ability of neurons to discriminate between images that are lateral and vertical mirror images of each other. Consequently, it does not abolish the overall advantage for vertical mirror images observed without training.

### 3.3.7 Fixation task: Trained vs. Untrained Orientations

In the previous section we observed that there was subtle, but significant increase in selectivity for trained compared to untrained images. We next asked whether these training effects generalized across all orientations of a trained image or were specific to only those orientations seen in discrimination training. Since the monkeys had been trained to discriminate only four of the eight orientations of the trained shapes (e.g., Figure 22, black box), the inclusion of untrained orientations of trained shapes (e.g., Figure 22, blue box) in the fixation task allowed us to address this question. Specifically, we asked whether the response rate, orientation selectivity and the degree to which images in lateral and vertical mirror image pairs were discriminated differed between trained and untrained orientations of trained shapes presented in the fixation task. Data were collected from 189 IT neurons (monkey Ph:  $n = 104$ ; monkey Op:  $n = 85$ ) during the passive presentation of trained and untrained orientations. Of these neurons, 168 were determined to be visually responsive (monkey Ph:  $n = 89$ ; monkey Op:  $n = 79$ ). Only these visually responsive neurons were considered in the following analyses. These

\* A loglinear analysis taking into account all three types of orientation pairs (lateral mirror image, vertical mirror image and 180° rotated pairs) revealed a main effect of training ( $p < 0.0001$ ) and a main effect of orientation ( $p < 0.0001$ ), but no interaction between training and orientation ( $p > 0.2$ ).

analyses proceeded in the same manner as for the comparison between trained and untrained shapes. Methods of analyses were identical except that in the previous comparisons, data from the two untrained tetrads were averaged together. In the present comparison, there is only one untrained tetrad, and therefore no averaging step was necessary.

#### Effects of Training on Firing Rate:

We first asked whether the overall response rate differed between trained and untrained orientations of trained shapes. The firing rate elicited by the preferred orientation in the trained tetrad was compared to the firing rate elicited by the preferred orientation in the untrained tetrad. For each visually responsive neuron, the mean response rate was calculated over the stimulus period for each orientation in the trained and untrained tetrads. The ‘preferred’ orientation was then identified as the orientation eliciting the strongest response. Statistical analyses were then performed on square root transformed data (see methods).

Figure 25 shows the mean firing rate for the preferred trained orientation plotted against the mean firing rate for the preferred untrained orientation for each neuron (monkey Ph: mean trained = 20.90 spikes/s, mean untrained = 20.12 spikes/s; monkey Op: mean trained = 29.11 spikes/s, mean untrained = 30.47 spikes/s). To assess whether the distributions differed significantly, we carried out a repeated measures ANOVA, with transformed response rate as the dependent variable, and training status (trained or untrained) and monkey (monkey Ph or monkey Op) as factors. This analysis revealed a significant difference in response rate between the two monkeys ( $p < 0.0005$ ), but no

significant main effect of training status ( $p > 0.6$ ), and no interaction between monkey and training status ( $p > 0.2$ ). We conclude that, as with the trained-untrained shape comparison, there is no difference in response rate for preferred trained and untrained orientations of trained shapes.

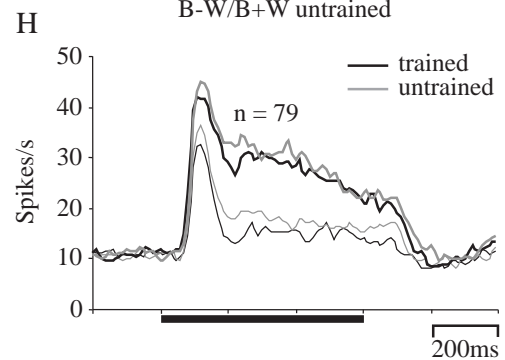
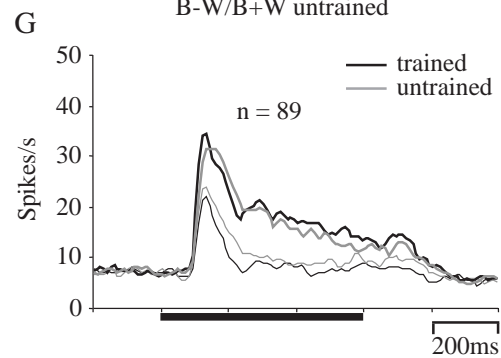
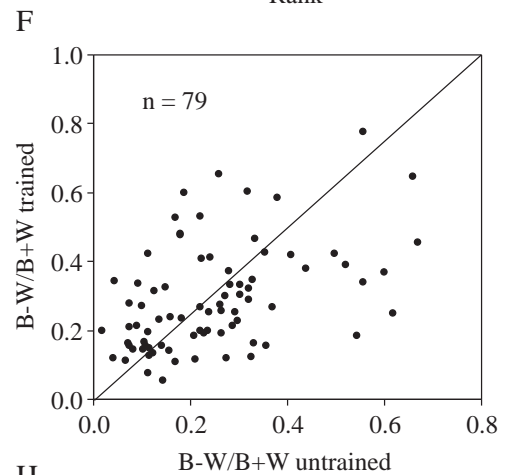
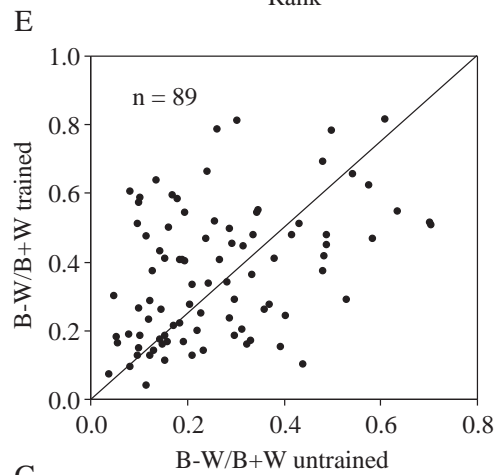
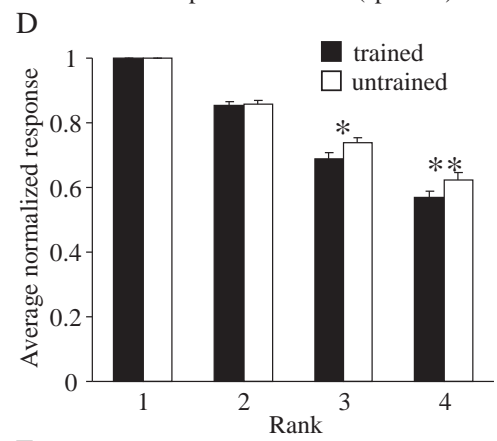
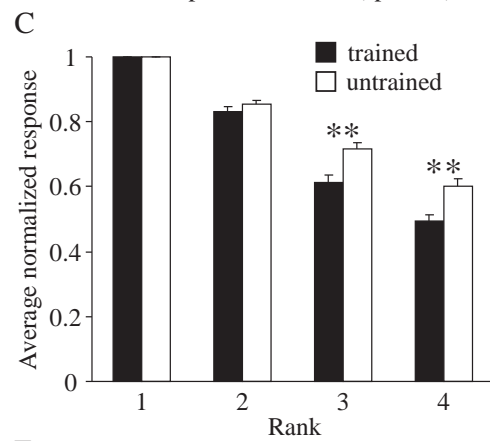
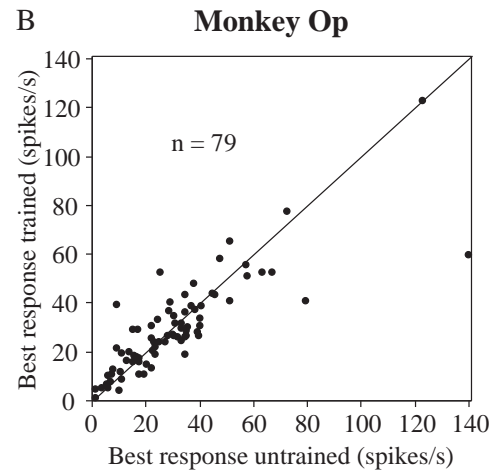
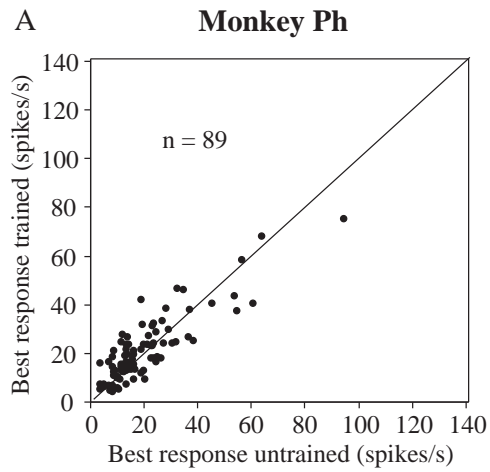
#### Effects of Training on Overall Selectivity:

We next asked if training increased selectivity for all orientations of trained shapes, or just those on which the monkey was trained. To answer this, we compared the orientation selectivity for trained and untrained orientations of trained shapes. Figure 25C-D shows the average rank responses to trained and untrained orientations normalized to the best responses for monkey Ph and Op, respectively. This analysis clearly shows

---

**Figure 25. Comparison of neural activity elicited by trained and untrained orientations of trained shapes presented in the fixation task.** A-B. Response to the best trained orientation plotted against the response to the best untrained orientation for monkey Ph (A) and monkey Op (B). Each point represents data from a single neuron. The ‘best’ orientation in each tetrad was defined as the one eliciting the strongest response. There was no significant tendency in either monkey for responses elicited by the best trained orientation to exceed those elicited by the best untrained orientation (ANOVA,  $p > 0.6$ ). C-D. Comparison of ranked neuronal responses elicited by trained orientations to ranked neuronal responses elicited by untrained orientations for monkey Ph (C) and monkey Op (D). Bars represent mean normalized response strengths across all neurons for trained orientations (black bars) and for untrained orientations (white bars). For each neuron, orientations in the trained and untrained tetrads were ranked from the most effective to the least effective. Firing rates were then normalized to the firing rate elicited by the best orientation in each tetrad. The mean firing rate at each rank of each tetrad was then computed across all sessions. Error bars represent standard error of the mean (s.e.m.). Asterisks indicate a significant difference between trained and untrained firing rates as determined by a post hoc analysis (Tukey HSD,  $*p < 0.05$ ,  $**p < 0.005$ ). A repeated measures ANOVA revealed a main effect of training ( $p < 0.00003$ ). E-F. Scatter plot of index values representing selectivity within tetrads of trained and untrained orientations for monkey Ph (E) and monkey Op (F). Index of selectivity =  $(b - w) / (b + w)$ , where  $b$  and  $w$  are response rates elicited by best and worst orientations in the tetrad. A repeated measures ANOVA revealed a main effect of training ( $p < 0.000002$ ) and a significant interaction between training and monkey ( $p < 0.05$ ). Post hoc analysis revealed that there was highly significant training effect in monkey Ph ( $p < 0.00001$ ) and a non-significant trend in monkey Op ( $p > 0.1$ ). G-H. Population histograms of neural activity collected during performance of the fixation task for monkey Ph (G) and monkey Op (H). Black lines represent activity during trials when trained orientations were presented and grey lines represent activity during trials when untrained orientations were presented. Thick lines represent responses for best orientations and thin lines represent responses for worst orientations. Activity was aligned on shape onset and black bars indicate duration of shape presentation. Tick marks on horizontal axis are in 200 ms increments. Histogram bin widths were 10 ms. Curves were smoothed according to the formula  $Y_n = 0.25Y_{n-1} + 0.5Y_n + 0.25Y_{n+1}$ , where  $Y_n$  is instantaneous firing rate.

---



that low-ranked trained orientations elicited weaker normalized responses than low-ranked untrained orientations, indicating that increases in selectivity are specific to only those orientations seen during training. To assess the significance of this effect, we carried out a repeated measures ANOVA, with normalized response strength as the dependent variable, and training status (trained or untrained orientation), rank (2-4) and monkey (monkey Ph or monkey Op) as factors. This analysis revealed an overall main effect of training status ( $p < 0.00003$ ), and while there existed a main effect of monkey ( $p < 0.05$ ), there was no interaction between monkey and training status ( $p > 0.1$ ). Additionally, there was a main effect of rank ( $p < 0.000001$ ) and a significant interaction between training status and rank ( $p < 0.000001$ ). *Post-hoc* analyses revealed that comparisons at ranks 3 and 4 were highly significant for both monkeys (monkey Ph: rank-3:  $p < 0.00002$ , rank-4:  $p < 0.00002$ ; monkey Op: rank-3:  $p < 0.03$ , rank-4:  $p < 0.004$ ).

We next asked whether the increase in selectivity for trained orientations was a result of low-ranked trained orientations eliciting significantly weaker response than low-ranked untrained orientations. Transformed response strengths at the lowest rank (rank-4) for trained and untrained orientations were compared using a repeated measures ANOVA, with transformed response strength as the dependent variable, and training status (trained or untrained orientation) and monkey (monkey Ph or monkey Op) as factors. This analysis revealed a main effect of training status ( $p < 0.00002$ ). That is, response rates were significantly less for rank-4 orientations of trained compared to untrained orientations (monkey Ph: mean trained = 10.69 spikes/s, mean untrained = 12.70 spikes/s; monkey Op: mean trained = 16.82 spikes/s, mean untrained = 19.28

spikes/s). While there was a main effect of monkey ( $p < 0.0005$ ), critically, there was no significant interaction between monkey and training status ( $p > 0.6$ ). These results suggest that the increase in selectivity for trained orientations over untrained orientations is a result of a decrease in response strength at low-rank orientations.

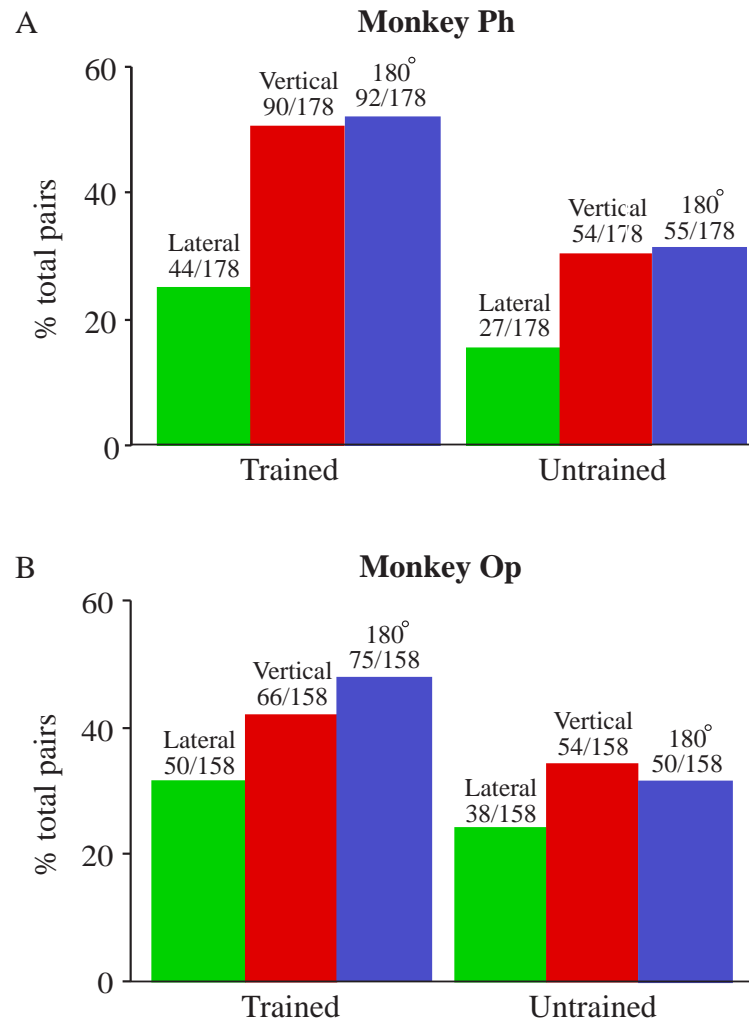
We next asked whether the increase in selectivity for trained orientations over untrained orientations was the product of large changes in a few neurons or small changes in many. To examine this, we calculated selectivity indices for the tetrad of trained orientations and the tetrad of untrained orientations for each neuron (see methods). These values are plotted separately for each monkey in Figure 25E-F. It appears that for monkey Ph, values were shifted toward a higher level of selectivity for trained orientations, and that this is due to a subtle shift in values across the population as a whole and not just in a subset of neurons. To assess the significance of this shift, we carried out a repeated measures ANOVA with selectivity index strength as the dependent variable, and training status (trained or untrained) and monkey (monkey Ph or monkey Op) as factors. Overall, there existed a main effect of training status ( $p < 0.000002$ ), a main effect of monkey ( $p < 0.04$ ), and a significant interaction between training status and monkey ( $p < 0.05$ ). *Post hoc* analyses revealed that there was a highly significant effect of training status in monkey Ph ( $p < 0.00001$ ), but no effect in monkey Op ( $p > 0.1$ ). The lack of a training effect in monkey Op stands in conflict with the results of the rank analysis described above. The rank analysis takes into account the neuronal responses to all four orientations and therefore may serve as a more sensitive measure of selectivity.



To determine the time course of the observed selectivity, we constructed population histograms of the best (rank-1) and worst (rank-4) orientations for both the trained and untrained tetrads. These histograms, constructed separately for the two monkeys, are shown in Figure 25G-H. Black lines indicate responses evoked by trained orientations and grey lines indicate responses evoked by untrained orientations. Thick and thin lines are responses to rank-1 (best) and rank-4 (worst) orientations, respectively. It is clear that for monkey Ph, there is greater selectivity for trained orientations than untrained orientations, and this increase in selectivity persists throughout the visual response. This is evident by the fact that there is a greater difference between response strength for best and worst trained orientations (black lines) than for best and worst untrained orientations (gray lines). For monkey Op, there appears to be little difference in selectivity for trained and untrained orientations. These results are in agreement with the results of the selectivity index analysis described above.

#### Effects of Training on Neuronal Discrimination Between Lateral and Vertical Mirror Images

It might be the case that training differentially increased the tendency of neurons to discriminate between lateral mirror images, thus altering the degree to which neurons exhibit lateral mirror image confusion. Therefore, we next sought to determine whether training the monkeys on the discrimination task differentially affected the discrimination of lateral and vertical mirror images. We counted instances in which neuronal activity discriminated between two images both for lateral and vertical mirror image pairs. This was done for the two lateral and two vertical mirror image pairs for the trained



**Figure 26. Frequencies of orientation pairs for trained and untrained orientations that elicited significantly different responses from neurons.** Data are shown separately for monkeys Ph (A) and Op (B). Frequencies are shown for lateral mirror image pairs (green bars), vertical mirror image pairs (red bars), and 180° rotations (blue bars). Values along ordinate represent the percent of orientation pairs that elicited significantly different responses out of all possible pairs. Significance was determined by a *t*-test ( $p < 0.05$ ).

orientations and for the two lateral and two vertical mirror image pairs for the untrained orientations. Figure 26A-B shows the frequencies of lateral pairs (green bars) and vertical pairs (red bars) that elicited significantly different responses for trained and untrained orientations in monkeys Ph and Op (frequencies of 180° rotated pairs that elicited significantly different responses are shown in blue for comparison). As for the comparison between trained and untrained shapes, training did not eliminate the tendency for neurons to discriminate between vertical mirror images more often than lateral mirror images. However, in accord with the analyses described above, the overall frequency of discrimination was increased with training. A loglinear analysis revealed that this was indeed the case. There was no main effect of monkey ( $X^2 = 2.25, p > 0.1$ ), but the main effect of training status ( $X^2 = 19.45, p < 0.001$ ) and image relation ( $X^2 = 36.33, p < 0.0001$ ) were significant. Critically, there was no interaction between training status and image relation ( $X^2 = 0.15, p > 0.7$ ), and no interaction between monkey and training status ( $X^2 = 2.39, p > 0.1$ )\*. We conclude that training to discriminate among the images in a tetrad enhances equally the ability of IT neurons to discriminate between images that are lateral and vertical mirror images of each other. Thus, it does not abolish the overall advantage for vertical mirror images observed without training.

### 3.3.8 DMS Task vs. Fixation task

It might be the case that the effects of discrimination training on the responses of IT neurons would differ depending on whether the monkey was performing the DMS task

\* A loglinear analysis taking into account all three types of orientation pairs (lateral mirror image, vertical mirror image and 180° rotated pairs) revealed a main effect of training ( $p < 0.0001$ ) and a main effect of orientation ( $p < 0.0001$ ), but no interaction between training and orientation ( $p > 0.6$ ).

or the fixation task. Therefore, to test whether the context in which images were presented influenced the neuronal responses in IT for those images, we compared overall response strength, degree of selectivity and degree to which images in lateral and vertical mirror image pairs were discriminated in the DMS and fixation tasks. For these comparisons, data collection procedures differed slightly between monkeys. In monkey Op, data collected in the DMS task were compared to data collected in the fixation task described in the methods section. Only data collected during the four trained orientation conditions were used in this analysis. In monkey Ph, data collected in the DMS task were compared to data collected during the presentation of the four trained orientations in a block of fixation trials separate from those collected for the trained-untrained comparisons described above. The task parameters for these trials were identical to those described in the methods section.

Data were collected from 168 IT neurons (monkey Ph:  $n = 85$ ; monkey Op:  $n = 83$ ) during the performance of the DMS task and the fixation task. Of these neurons, 152 were determined to be visually responsive in at least one condition in at least one task (monkey Ph:  $n = 75$ ; monkey Op:  $n = 77$ ). Only these visually responsive neurons were considered further. During data collection, the order in which the blocks were run varied from session to session. For monkey Ph, 32 neurons were run on the fixation task first and 43 were run on the DMS task first. For monkey Op, 28 neurons were run on the fixation task first and 49 were run on the DMS task first.

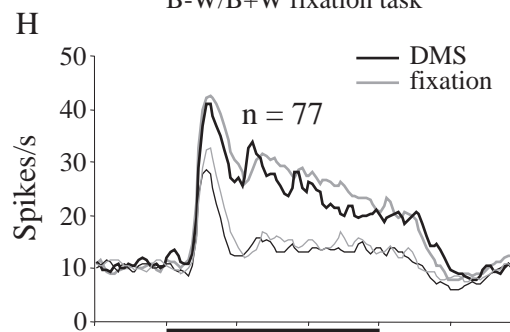
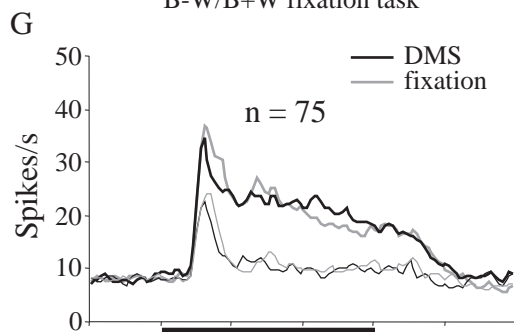
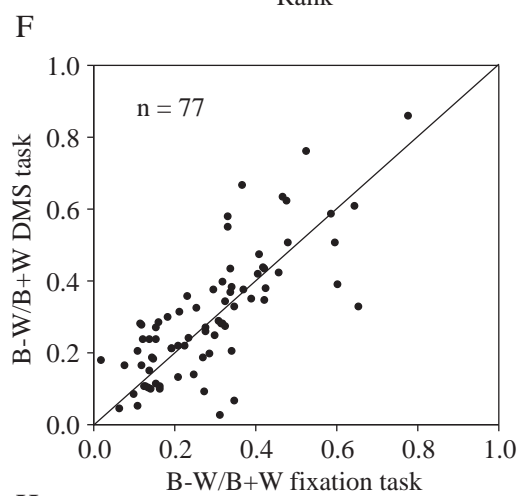
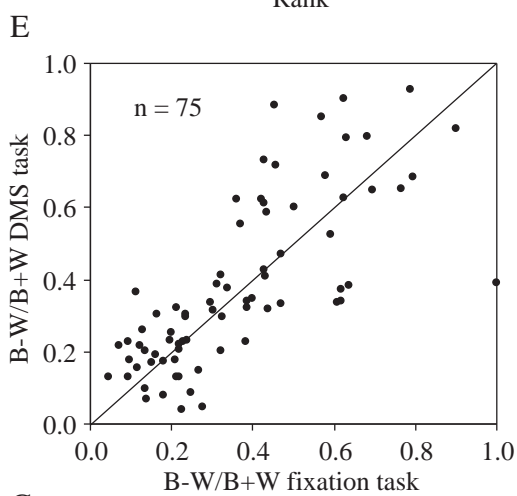
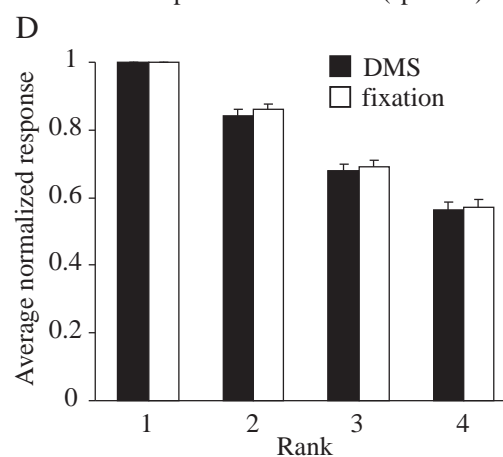
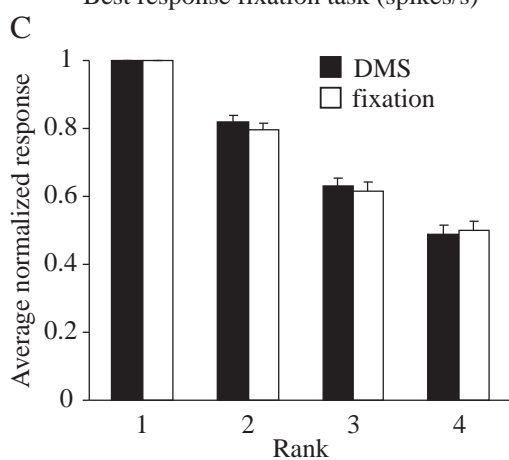
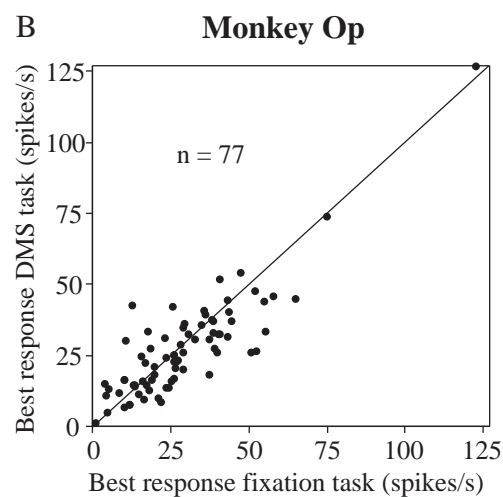
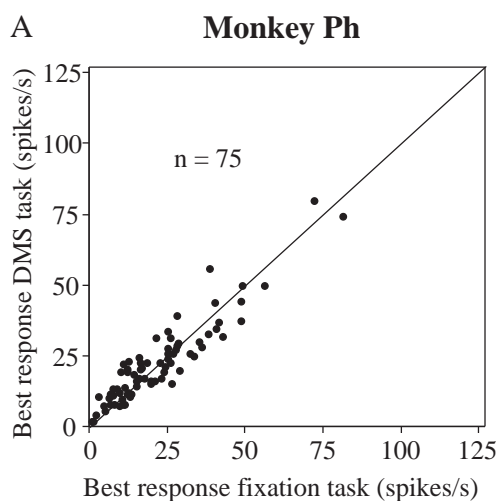
### Effects of Task Context on Firing Rate

To determine whether the overall response strength for trained orientations was affected by task context, we compared the firing rate elicited by the preferred trained orientation in the fixation task to the firing rate elicited by the preferred trained orientation in the DMS task. The “preferred” orientation was determined separately for each task. Figure 27A-B shows the mean firing rate for the preferred trained orientation in the DMS task plotted against the mean firing rate for the preferred trained orientation in the fixation task for each neuron (monkey Ph: mean DMS = 22.23 spikes/s, mean fixation = 22.55 spikes/s; monkey Op: mean DMS = 26.96 spikes/s, mean fixation = 29.35 spikes/s). To assess whether the distributions differed significantly, a repeated measures ANOVA was performed with transformed response rate as the dependent

---

**Figure 27. Comparison of neural activity elicited by trained orientations presented in the fixation task and DMS task.** A-B. Response to the best trained orientation in the DMS task plotted against the response to the best trained orientation in the fixation task for monkey Ph (A) and monkey Op (B). Each point represents data from a single neuron. The ‘best’ orientation in each task was defined as the one eliciting the strongest response. There was no significant difference in response strength evoked by the best trained orientations in the two tasks (ANOVA,  $p > 0.1$ ). C-D. Comparison of ranked neuronal responses elicited by trained orientations in the two tasks for monkey Ph (C) and monkey Op (D). Bars represent mean normalized response strengths across all neurons for trained orientations in the DMS task (black bars) and fixation task (white bars). For each neuron, trained orientations in the two tasks were ranked from the most effective to the least effective. Firing rates were then normalized to the firing rate elicited by the best orientation in each task. The mean firing rate at each rank of each tetrad was then computed across all sessions. Error bars represent standard error of the mean (s.e.m.). A repeated measures ANOVA revealed that there was no significant effect of task ( $p > 0.8$ ). E-F. Scatter plot of index values representing selectivity for trained orientations within the two tasks for monkey Ph (E) and monkey Op (F). Index of selectivity =  $(b - w) / (b + w)$ , where  $b$  and  $w$  are response rates elicited by best and worst orientations in the tetrad. The distribution of index values across all neurons was not significantly different for the two tasks ( $p > 0.3$ ). G-H. Population histograms of neural activity collected during performance of the DMS and fixation tasks for monkey Ph (G) and monkey Op (H). Black lines represent activity evoked by trained orientations during the DMS task and grey lines represent activity evoked by trained orientations during the fixation task. Thick lines represent responses for best orientations and thin lines represent responses for worst orientations. Activity was aligned on shape onset and black bars indicate duration of shape presentation. Tick marks on horizontal axis are in 200 ms increments. Histogram bin widths were 10 ms. Curves were smoothed according to the formula  $Y_n' = 0.25Y_{n-1} + 0.5Y_n + 0.25Y_{n+1}$ , where  $Y_n$  is instantaneous firing rate.

---



variable, and task context (DMS or fixation task), monkey (monkey Ph or monkey Op) and block order (DMS run first or fixation run first) as factors. While this analysis revealed a significant difference in firing rate between the two monkeys ( $p < 0.006$ ), there was no significant main effect of task ( $p > 0.1$ ), no effect of block order ( $p > 0.3$ ) and no interaction between task and block order ( $p > 0.6$ ) or task and monkey ( $p > 0.09$ ). Therefore, we conclude that response strength of neurons is not influenced by task context.

#### Effects of Task Context on Overall Selectivity

We next asked whether task context influenced the selectivity of neurons for trained orientations. To answer this, we first ranked responses elicited by the four trained orientations separately for responses recorded in the DMS task and the fixation task. Then, we then normalized the responses elicited by each of the three low ranked orientations to the response elicited by the best orientation in each task. The normalized responses for the four ranks averaged across all neurons are shown in Figure 27C-D. There appears to be no difference in selectivity for trained orientations in the two tasks for either monkey, indicating that task context did not influence the selectivity of IT neurons for trained orientations. To confirm this observation, we carried out a repeated measures ANOVA, with normalized response strength as the dependent variable, and task (DMS or fixation), rank (2-4), monkey (monkey Ph or monkey Op) and order (DMS run first or fixation run first) as factors. This analysis revealed no overall main effect of task ( $p > 0.8$ ), a main effect of rank ( $p < 0.000001$ ), but no interaction between task and rank ( $p > 0.7$ ). There was a main effect of monkey ( $p < 0.03$ ), but no interaction between

monkey and task ( $p > 0.1$ ). Finally, there was no effect of order ( $p > 0.4$ ) and no interaction between task and block order ( $p > 0.4$ ).

In order to confirm that the absence of selectivity differences between the two tasks was consistent across all neurons, we calculated a selectivity index for the tetrad of trained orientations presented in the DMS and fixation tasks, for each neuron (see methods). These values are plotted in Figure 27E-F. We found that across all neurons, values were not significantly different for trained orientations presented in either task. A repeated measures ANOVA with selectivity index strength as the dependent variable, and task (DMS or fixation) and monkey (monkey Ph or monkey Op) as factors confirmed this observation. There was no main effect of task context ( $p > 0.3$ ), and while there was a main effect of monkey ( $p < 0.02$ ), there was no significant interaction between monkey and task context ( $p > 0.9$ ). In addition, there was no main effect of block order ( $p > 0.6$ ), and no interaction between task context and block order ( $p > 0.06$ ).

Since data used in the above analyses were firing rates averaged over a large portion of the stimulus response, it might be the case that selectivity varied as a function of time and that this was not accounted for in the above analyses. Therefore, we constructed population histograms of the best (rank-1) and worst (rank-4) orientations for the trained orientations run in the DMS and fixation tasks. These histograms, constructed separately for the two monkeys, are shown in Figure 27G-H. Black lines indicate responses evoked by trained orientations and grey lines indicate responses evoked by untrained orientations. Thick and thin lines are responses to rank-1 (best) and rank-4 (worst) orientations, respectively. It appears that selectivity for trained and untrained orientations varied little over the response period. Moreover, there was little difference

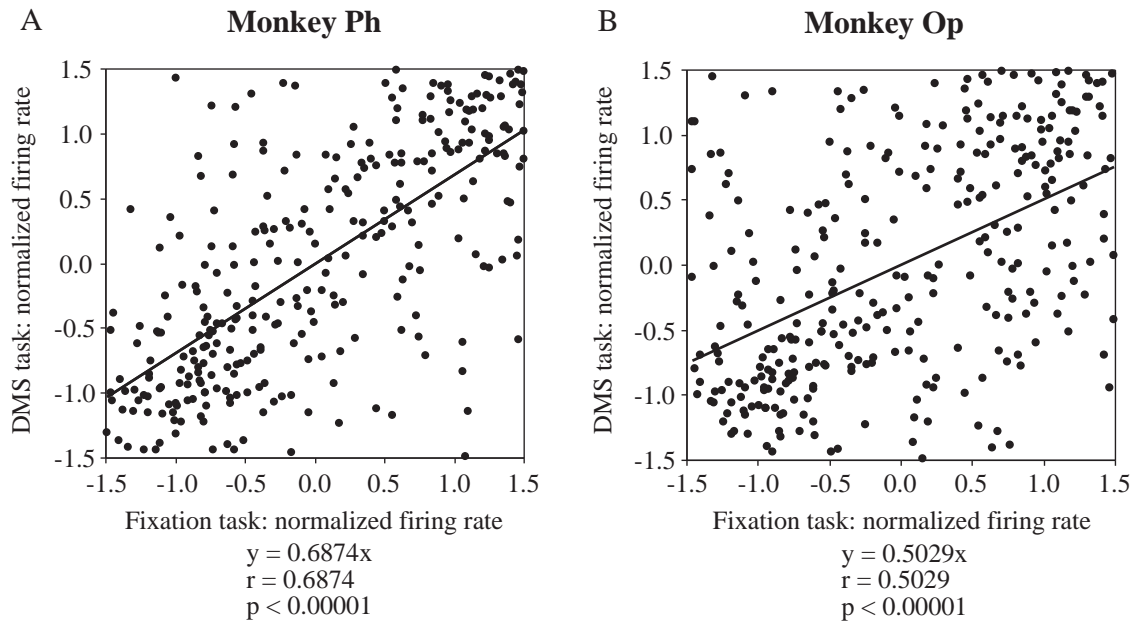


in selectivity for trained and untrained orientations, confirming the results of the above analyses.

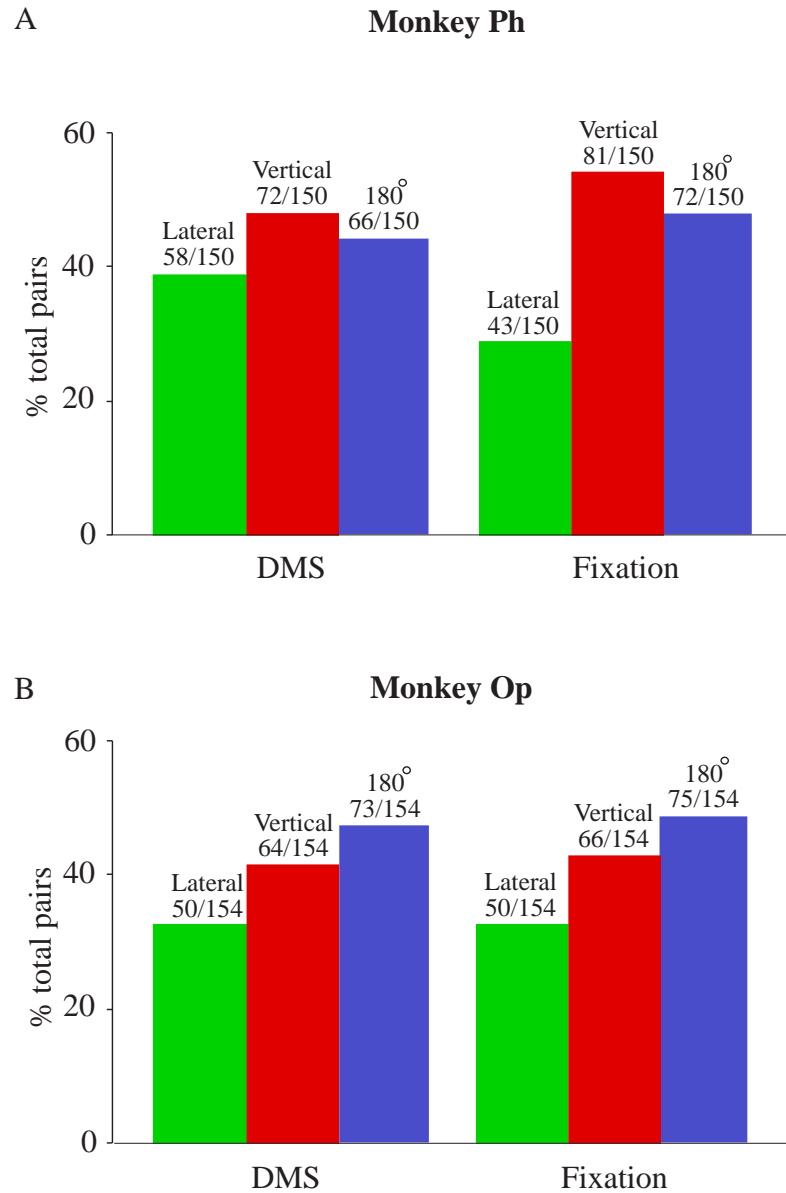
While the previous analyses addressed the degree to which neurons exhibited orientation selectivity in the two tasks, they did not evaluate whether the specific orientation preferences in the two tasks differed. We therefore asked whether there was a correlation between orientation preferences of neurons for trained orientations presented in the DMS and fixation tasks. To test this, we determined, for each neuron, the firing rates elicited by the four trained orientations in the two tasks. We then normalized each of the four firing rates in a given task according to the formula  $X' = (X - a) / std$ , where  $X$  is the average firing rate for a given orientation,  $a$  is the average of the four firing rates and  $std$  is the standard deviation of the four firing rates. The normalized firing rates in the two tasks for each orientation of each neuron are plotted in Figure 28A-B for the two monkeys. A correlation analysis revealed a highly significant correlation between the selectivity for trained orientations in the two tasks in each monkey (Pearson product-moment correlation: monkey Ph:  $r = 0.6874$ ,  $p < 0.000001$ ; monkey Op:  $r = 0.5029$ ,  $p < 0.000001$ ). Therefore, we conclude that overall, the selectivity for specific orientations in the two tasks was relatively unchanged.

#### Effects of Task Context on Neuronal Discrimination Between Lateral and Vertical Mirror Images

Finally, we asked whether task context affected the degree to which IT neurons tend to confuse lateral mirror images more than vertical mirror images. Figure 29A-B shows the frequencies of lateral pairs (green bars) and vertical pairs (red bars) that



**Figure 28. Correlation between neuronal responses for trained orientations presented in the DMS and fixation tasks.** Data are shown separately for monkey Ph (A) and monkey Op (B). Each point represents the normalized firing rate of a single neuron elicited by a trained orientation presented in the two tasks. Firing rates were normalized according to the formula  $\bar{X}' = (X - a) / std$ , where  $X$  is the average firing rate for a given orientation,  $a$  is the average, and  $std$  is the standard deviation of the four firing rates evoked by the four trained orientations. Values on abscissa are normalized firing rates evoked by trained orientations presented in the fixation task. Values on the ordinate are normalized firing rates evoked by trained orientations presented in the DMS task. Regression lines are superimposed over the data. There was a significant correlation in both monkey Ph ( $r = 0.6874$ ;  $p < 0.00001$ ) and monkey Op ( $r = 0.5029$ ;  $p < 0.00001$ ).



**Figure 29. Frequencies of orientation pairs for trained orientations run in the DMS and fixation tasks that elicited significantly different responses from neurons.** Data are shown separately for monkeys Ph (A) and Op (B). Frequencies are shown for lateral mirror image pairs (green bars) and vertical mirror image pairs (red bars). Values along ordinate represent the percent of orientation pairs that elicited significantly different responses out of all possible pairs. Significance was determined by a *t*-test ( $p < 0.05$ ).

elicited significantly different responses in the two tasks for monkeys Ph and Op (frequencies of 180° rotated pairs that elicited significantly different responses are shown in blue for comparison). In both tasks, neurons were selective for more vertical mirror image pairs than lateral mirror image pairs. However, there appeared to be no consistent effect of task context on the number of discriminated pairs in either category. A loglinear analysis confirmed this observation. There was no main effect of monkey ( $X^2 = 2.83, p > 0.09$ ), no main effect of task ( $X^2 = 0.11, p > 0.7$ ), but a significant main effect of image relation ( $X^2 = 23.12, p < 0.0001$ ). Critically, there was no interaction between task and image relation ( $X^2 = 2.44, p > 0.1$ ) and no interaction between monkey and task ( $X^2 = 0.31, p > 0.5$ )\*. We conclude from this analysis that neurons exhibit lateral mirror image confusion regardless of task context. These results also confirm that task context does not influence the degree of orientation selectivity of IT neurons.

### 3.4 Discussion

#### 3.4.1 Overview

We recorded single-neuron activity in the inferotemporal cortex of two monkeys trained to discriminate among different orientations of each of ten shapes. The results lead to several conclusions: (1) training did not result in a change in response strength for preferred orientations, (2) training significantly enhanced neuronal selectivity among trained orientations, (3) this increase in selectivity took the form of small changes in

\* A loglinear analysis taking into account all three types of orientation pairs (lateral mirror image, vertical mirror image and 180° rotated pairs) revealed no main effect of task ( $p > 0.9$ ), a main effect of orientation ( $p < 0.0001$ ), and no interaction between task and orientation ( $p > 0.2$ ).

many neurons, (4) training to discriminate between both lateral and vertical mirror images did not alter the tendency for neurons to discriminate between vertical mirror images more often than lateral mirror images, (5) neuronal selectivity for trained orientations recorded in the DMS task was significantly correlated with the monkeys' ability to discriminate those orientations, and (6) selectivity for trained orientations was not affected by task context. In the following sections, we consider these findings in relation to the existing literature on properties of IT neurons and the impact of discrimination training on these properties.

#### *3.4.2 Effects of Training on Neuronal Response Strength*

Overall, we found no significant effect of orientation discrimination training on the net strength of responses elicited by trained orientations. This was true for comparisons of trained orientations to both untrained shapes and untrained orientations of trained shapes. This finding stands in conflict with several studies demonstrating that response strength increases with discrimination training. These discrepancies may be accounted for by differences in experimental design. For example, Kobatake et al. (1998) observed an increase in the response strength of neurons as a result of discrimination training. However, comparisons were made between trained and untrained monkeys and in one case between two hemispheres of a single animal before and after training. Between-neuron comparisons are more subject to noise than within-neurons comparisons. Thus, without conducting within-neuron comparisons of trained and untrained stimuli, one cannot fully conclude that these observed differences in response strength are purely experience based. Two other studies also demonstrated training effects on response

strength of IT neurons (Miyashita et al., 1993; Sakai and Miyashita, 1994). In both studies, however, neurons were chosen for study on the basis of their responsiveness to trained stimuli. This selection procedure could well have biased the sample toward neurons more responsive to trained than to untrained stimuli.

Recent experiments in our laboratory have suggested that training does not influence the overall response strength of neurons in IT. Baker et al. (2002) trained monkeys on a feature conjunction task and demonstrated that there was no difference in responses to preferred trained and untrained stimuli. However, this study used unusual baton-like stimuli with top and bottom features to which the monkey was required to attend. It is not clear whether the absence of training effects on response strength was a result of such unique stimuli and task requirements. In the present study, we demonstrate that this finding is not specific to the methods used by Baker et al. We found that training on an orientation discrimination task does not result in changes in overall response strength of neurons for trained images.

#### *3.4.3 Effects of Training on Neuronal Selectivity for Shape Orientation*

Our results demonstrate that orientation discrimination training significantly increases the selectivity of IT neurons for shape orientation. This is in disagreement with several studies that have failed to find an effect of training on neuronal selectivity. In one study, Vogels and Orban (1994b) trained monkeys to perform a successive orientation discrimination task with line gratings and then recorded from neurons while monkeys performed the task with trained and untrained orientations. The authors failed to find a significant effect of training on the selectivity of IT neurons. One reason for the

discrepancy between their study and ours may be that the line gratings may not have sufficiently driven IT neurons, as IT neurons typically respond best to more complex stimuli (Tanaka et al., 1991). Further, as suggested by the authors, the discrimination of the orientation of simple lines and gratings may not require processes within IT and may in fact be resolved at early stages within the ventral stream. A recent study by Erickson et al. (2000) also failed to find an effect of training on the selectivity of IT neurons. However, in this study, monkeys were only trained on stimuli for one day prior to recording, which may not have been enough to result in significant changes in neuronal selectivity.

In contrast to these studies, Logothetis and colleagues claimed to observe increases in selectivity due to discrimination training. However, these authors failed to provide a quantitative analysis of population responses to support this claim (Logothetis and Pauls, 1995; Logothetis et al., 1995). Two other groups have demonstrated effects of training on the selectivity of IT neurons. Kobatake et al. (1998) found that training increased selectivity of IT neurons, but made comparisons between trained and untrained animals. Thus, any effects observed may have been due to inter-animal differences. Finally, Baker et al. (2002) recorded responses of neurons in monkeys trained to perform a feature conjunction task and found that selectivity was greater for trained than untrained baton-like stimuli. In the present study, we have generalized this finding to orientation discrimination training. We demonstrated that training to discriminate orientation increases selectivity for trained orientations relative to untrained images. This was true not only for trained vs. untrained shapes, but also for trained vs. untrained orientations.

#### *3.4.4 Extent of Selectivity Changes Across Neuronal Populations*

We examined orientation selectivity for trained and untrained stimuli across the population of neurons studied and found that there was a general shift in selectivity over the entire population of neurons as opposed to a strong shift in a small subset of neurons. These results are not concordant with a report by Logothetis et al. (1995), which suggests that experience-based changes in the selectivity of IT neurons are the result of marked increases in selectivity in small number of neurons (Logothetis and Pauls, 1995; Logothetis et al., 1995). However, these authors do not offer quantitative analyses to support this claim. Our results are in accord with those found previously in this laboratory (Baker et al., 2002). Further, they are consistent with the findings of Kobatake et al. (1998), which showed that selectivity for trained shapes was broadly tuned in neurons and that this broad tuning lead to greater selectivity across neurons in trained monkeys compared to untrained monkeys.

#### *3.4.5 Lateral Mirror Image Confusion in IT Neurons*

We demonstrated in Chapter 2 that neurons in IT discriminate less effectively between lateral than between vertical mirror images. We have hypothesized that this functional trait may underlie lateral mirror image confusion as demonstrated in psychological and comparative behavioral studies (Sutherland, 1960; Rudel and Teuber, 1963; Riopelle et al., 1964; Huttenlocher, 1967; Sekuler and Houlihan, 1968; Serpell, 1971; Hamilton and Tieman, 1973; Bornstein et al., 1978; Todrin and Blough, 1983). One possible cause of this difference in selectivity is that organisms are not commonly required to discriminate between lateral mirror images, as they convey little information



about the object. Thus, it would not be necessary for IT neurons to discriminate between them. If this is true, then we might expect that if the animal were required to discriminate between lateral mirror images, then IT neurons would become better at discriminating between them. We tested this in the present study by training monkeys to discriminate between lateral mirror image pairs, vertical mirror image pairs, and 180° in-plane rotations and then measuring the degree to which training influenced the selectivity of IT neurons for those three types of pairs. We found that training monkeys to discriminate between both lateral and vertical mirror image pairs resulted in a general increase in neuronal selectivity for both types of reflections. It did not affect the relative selectivity of neurons for lateral and vertical mirror images. It remains a question whether training confined to lateral mirror image pairs would result in the selective increase in the ability of neurons to make lateral mirror image discriminations

#### *3.4.6 Behavior in Relation to Neuronal Activity During Task Performance*

Relating neuronal activity to perceptual report is important if we are to understand the functional implications of the neural responses observed in experiments. In the present study, we considered this by recording the activity of IT neurons while the monkey performed the DMS task. We demonstrated that there is significant positive correlation between how well IT neurons discriminate between shape orientations and how well the monkey discriminates between them.

Few studies have examined the correlation between perception of objects and neural activity in IT. In one, Messinger et al. (2001) examined the responses of IT neurons over a single session in which monkeys learned to associate pairs of stimuli.

They demonstrated that the ability of IT neurons to encode paired stimuli paralleled the degree to which monkeys learned those associations. In another study, Op de Beeck et al. (2001) demonstrated that the degree to which monkeys discriminated parameterized stimuli was congruent with the ability of IT neurons to discriminate them. The results of these two studies demonstrate that the ability to associate and discriminate complex shapes is reflected in the activity of IT neurons.

In the present study, we consider an issue not previously addressed, and that is whether the activity of IT neurons reflects the ability of animals to discriminate shape orientation. Ours is the first study demonstrating that the selectivity of IT neurons for shape orientation parallels the behavioral ability to discriminate of shape orientation. This finding raises the question of whether activity within IT contributes to the process of orientation discrimination. Support for this idea comes from a study by Gross (1978) which examines the effects of IT lesions on the ability of monkeys to discriminate shapes and shape orientations. The results show that lesions of IT result in greater impairment in shape discrimination than in orientation discrimination. However, the data demonstrate that IT lesioned animals were, in fact, impaired at orientation discrimination relative to normals. These data, along with our findings, suggest that IT may play some role in the discrimination of shape orientation.

#### *3.4.7 Effects of Task Context on Neuronal Responses*

We have demonstrated that presenting images from the training set in a DMS task had little effect on the response strength or selectivity of IT neurons compared to that in a fixation task. Although the timing in the fixation task was similar to that of the initial

portion of the DMS task, the two trial types were presented in blocks. This argues against the possibility that the observed result is due to the monkeys covertly performing a memory task in the fixation task.

Although some studies have noted differences between the delay period activity following sample offset in a DMS task and the activity in the period following stimulus offset in a fixation task (Vogels et al., 1995; Chelazzi et al., 1998), few have found effects of task context on visual responses in IT. For example, Lehky and Tanaka (2002) found that there were no changes in stimulus selectivity when stimuli were presented in either a passive viewing or a memory intensive task. Although the authors observed a decrease in response strength for stimuli presented in the fixation task, this effect was weak (only a 5% decrease in response strength for TE neurons). Further, Baker et al. (2002) reported that neither the response magnitude nor the degree of selectivity of IT neurons for baton-like stimuli differed between presentation in a fixation or feature conjunction task. Our results are consistent with those of Lehky and Tanaka (2002) and Baker et al. (2002), demonstrating that task context has little effect on the selectivity and response magnitude of IT neurons.

#### *3.4.8 Delay Period Activity*

Differential activity in IT neurons during the delay following sample offset in memory tasks is thought to reflect the identity of the sample that it to be remembered (Miyashita and Chang, 1988). Thus, it is hypothesized that this activity is involved in short term memory processes. In the present study, we found very few neurons that exhibited selective delay activity. This stands in conflict with several studies that have

demonstrated differential delay activity (Fuster and Jervey, 1981; Miyashita and Chang, 1988; Fuster, 1990; Miller et al., 1993b; Chelazzi et al., 1993; Vogels and Orban, 1994a; Miller and Desimone, 1994; Vogels et al., 1995; Gibson and Maunsell, 1997; Chelazzi et al., 1998). For example, Miyashita and Chang (1988) observed that approximately half of the task-related neurons studied exhibited selective delay period activity. Miller et al. (1993b) also found a large proportion of neurons (25% of visual neurons) exhibiting selective delay activity in a DMS task. On the other hand, our findings are consistent with those of Baylis and Rolls (1987), who reported finding only 2 neurons (out of 94 neurons) that exhibited selective delay activity in a DMS task.

Critically, many studies reporting differential delay activity have failed to examine (or report) whether the pattern of preferences during the delay period is the same as that in the sample period. (Miyashita and Chang, 1988; Fuster 1990; Gibson and Maunsell, 1997). Those groups that have compared the selectivity in these two periods have found varying results. For example, Fuster and Jervey (1981) reported that in the few neurons that exhibited delay activity (5-10%), only a subset of these had selectivity in the delay period that matched the sample. In contrast to this, Miller et al. (1993b) found that the average delay activity following the “best” sample was significantly greater than that following the “worst” sample for 37 cells that showed differential delay activity, suggesting that the selectivity of the delay activity was similar to that for the sample. In the present study, we failed to find any evidence for orientation preference in the delay period matching that of the sample period.

The differences in reported delay activity cited above may be due to differences in task requirements and recording sites. In the present experiment, the length of the delay

period was substantially shorter than that of previous studies, which often exceeded 1.5 seconds (Chelazzi et al. 1998; Gibson and Maunsell, 1997; Baylis and Rolls, 1987) and in some studies extended over 10 seconds (Miyashita and Chang, 1988; Fuster and Jervey, 1981; Fuster, 1990). While our recording sites were confined to the lower bank of the STS and the ventral convexity, some groups recorded in the more medial perirhinal cortex (Miller et al., 1993b). Recent studies have suggested that mnemonic encoding is greater in perirhinal cortex than in TE (Naya et al., 2003), which may have resulted in greater occurrence of delay activity.

Thus, the role of delay activity in IT in mnemonic encoding of stimuli is not clear. We conclude from the results of our study that in our task, the activity of IT neurons is most likely involved in the processing of the stimuli but not holding them in working memory.

#### *3.4.9 Behavioral Confusion of Lateral Mirror Images*

In analyzing the behavioral error patterns of the monkeys during recording sessions, we observed that the degree to which the monkeys exhibited lateral mirror image confusion seemed to be dependent on the particular shape being discriminated. That is, for some shapes, monkeys made primarily lateral mirror image errors and for others, primarily vertical mirror image errors. Only one image resulted in primarily 180° rotation errors (see Table 3). This stands in conflict with the studies demonstrating preponderant lateral mirror image confusion in monkeys regardless of the particular stimulus (Riopelle et al., 1964; Hamilton and Tieman, 1973). One possible reason for the discrepancy between our results and these is that our stimuli were more complex and our

task more demanding. In addition, the monkeys in our study were highly overtrained whereas there was no training required in the other studies. For example, Riopelle et al. (1964) initially trained monkeys to perform a simultaneous discrimination with color stimuli. When the animals were proficient at performing the task, the stimuli were switched to mirror image and rotated line patterns. The degree to which monkeys exhibited lateral mirror image confusion was determined by calculating the number of trials performed to reach criterion for these stimuli. In contrast, the monkeys in our study were trained to discriminate mirror image stimuli for several months. It could be that overtraining on a difficult task led monkeys to adopt elaborate strategies to solve the task. One way monkeys may have solved the task would be to use a template matching strategy. Alternatively, they may have employed a spatial strategy. For example, animals could identify a particularly salient feature on the image, identify the location of that feature, and then during the search period look for the image with that feature at that location. It could be, then, that the behavioral data collected in the present task conditions do not accurately reflect the perception of mirror images. It would be of interest to determine whether monkeys never trained on an orientation discriminate task would exhibit lateral mirror image confusion with our stimuli.

### **3.5 Summary**

In conclusion, we have shown that training monkeys to discriminate among different orientations of a shape increases the ability of neurons in IT to discriminate among the trained orientations. Moreover, we have observed a correlation between neuronal selectivity for orientation and the behavioral discrimination of orientation,

suggesting that the perception of shape orientation is reflected in the selectivity of these neurons. Finally, we have observed that orientation discrimination training, although enhancing selectivity overall, left intact the tendency of IT neurons to confuse lateral mirror images more often than vertical. These findings suggest that experience based changes in perception are supported by mechanisms within IT, and implicate IT in processes underlying shape orientation discrimination.

## Chapter 4

### Low frequency oscillations arising from competitive interactions between visual stimuli

#### 4.1 Introduction

In the previous two chapters, we examined the responses of IT neurons to complex images, as determined by the average response rate calculated over several hundreds of milliseconds. To a first approximation, information about such images carried by neuronal activity in IT is in the form of a rate code: the spike count throughout the response period is greater in the presence of a preferred image than of a non-preferred image. However, the temporal pattern of a neuronal response, as distinct from its net strength, can also vary depending on the image (Richmond et al., 1987; Richmond and Optican, 1987; Optican and Richmond, 1987). In this chapter, we investigate a specific temporal pattern observed in the activity of IT neurons. This is the tendency of IT neurons to respond to some stimuli with a series of discrete bursts of action potentials. Neuronal visual responses in the form of damped initially positive oscillations at a frequency of around 5 Hz were first described in studies of the temporal pole by Nakamura et al. (1991, 1992). Although not a major focus of other studies, such responses are evident in several published post-stimulus-time histograms. Two clear examples are contained in Figure 5A of Sheinberg and Logothetis (1997). Less dramatic but still convincing instances of oscillation, in the form of a peak, followed by a trough, followed by rebound, appear in histograms accompanying several other papers (Sato et al., 1980, Figure 3A-B; Sato, 1989, Figure 3E; Tamura and Tanaka, 2001, Figures 3-4).



The functional significance of this low-frequency oscillatory activity is not known. Nakamura et al. (1991) reported that oscillatory responses were elicited more often by familiar than by unfamiliar stimuli. However, 5 Hz oscillations have not been cited as a distinguishing feature in studies comparing responses elicited by learned or familiar stimuli to responses elicited by unlearned or novel ones (Logothetis and Pauls, 1995; Booth and Rolls, 1998; Kobatake et al., 1998; Xiang and Brown, 1998, 1999; Erickson and Desimone, 1999; Erickson et al., 2000; Baker et al., 2002).

In the present study, we investigated the hypothesis that oscillatory activity in IT arises from competition among neurons selective for different images. Networks in which neurons compete with each other and are subject to fatigue can alternate between states in which antagonistic populations are active (Wilson et al., 2000). Oscillatory visual responses therefore might occur in IT through a process in which neurons responsive to an image are initially excited by it, then succumb through fatigue to suppression by other neurons not responsive to the image, then recover and fire again, and so on. Intrinsic inhibitory connections mediated by GABAergic interneurons, a well known feature of IT (Wang et al., 2000), may play a role in the competitive interactions among stimuli that have been described by several groups (Chelazzi et al., 1993, 1998; Desimone, 1998; Miller et al., 1993a; Missal et al., 1999; Moran and Desimone, 1985; Sato, 1989, 1995). The idea that IT neurons are subject to fatigue is compatible with the fact that the activity of excitatory cortical neurons adapts to prolonged electrical stimulation (Connors and Gutnick, 1990) and with the observation that IT neurons commonly respond to prolonged visual stimulation with a phasic burst that tapers off to a tonic plateau (Tamura and Tanaka, 2001).

If oscillatory visual responses in IT depend on competition between neurons selective for different patterns, then the tendency for an excitatory stimulus to elicit an oscillatory response from a neuron should be increased by the presence of other images - images to which the neuron is not responsive but to which its competitors do respond. The aim of the experiments described here was to test that hypothesis and thus to elucidate the nature of mechanisms underlying oscillatory visual responses in IT.

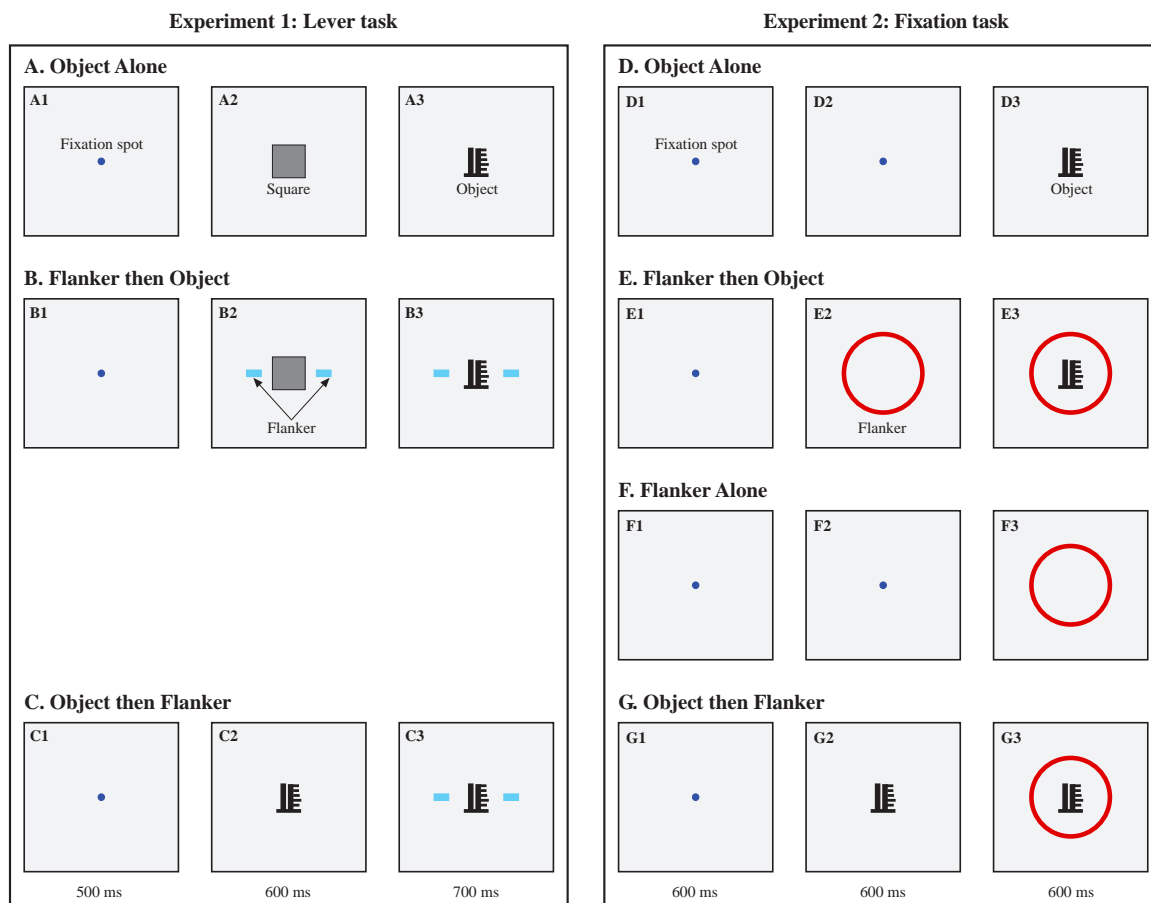
## 4.2 Methods

### 4.2.1 Subjects

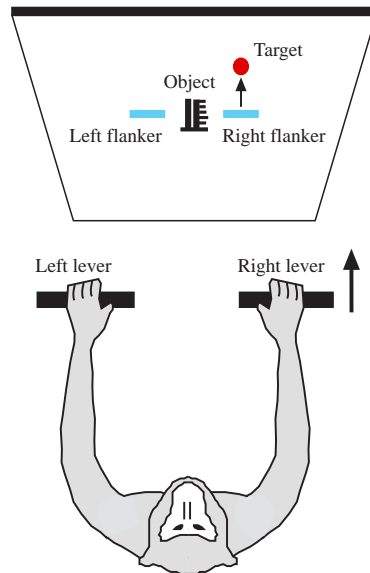
Three adult male rhesus monkeys (*Macaca mulatta*) weighing 5.5-7.5 kg were used in this experiment. Their laboratory designations were Op, Fi and Ph. General surgical and training procedures are described in Appendix A. Training specific to this experiment is described below.

### 4.2.2 Stimuli

Stimuli were selected from a library of 107 moderately complex white shapes, approximately  $3^\circ \times 3^\circ$  in height and width. Examples of these shapes are shown in Chapter 2, Figure 4A. Prior to collecting data from a neuron, shapes were presented sequentially to the fixating monkey. We selected one shape which elicited strong excitatory responses, hereafter referred to as the 'object', for further study of each neuron.



**Figure 30. Data collection tasks.** Sequence of visual stimuli under three conditions in Experiment 1 (A-C) and four conditions in Experiment 2 (D-G). Throughout each sequence, the monkey maintained central fixation. The time at the bottom of each column indicates the duration of the display depicted in that column. The centrally presented object was selected to elicit excitatory responses from the recorded neuron and so varied from session to session.



**Figure 31. Lever response in Experiment 1.** In Experiment 1, subsequent to the sequence of visual stimuli depicted in Figure 30A-C, a target appeared on the screen and the monkey made a lever response contingent on the target's location. While the monkey (seated upright and drawn as seen from above) fixated the center of an upright screen and grasped two levers, the target appeared at one of four locations relative to the center of the screen: (1) up and to the left, (2) up and to the right, (3) down and to the left or (4) down and to the right. To the four targets, respectively, the monkey had to respond with the following actions on the levers: (1) left lever forward, (2) right lever forward, (3) left lever back, or (4) right lever back. Whichever lever the monkey moved, the corresponding flanker, if visible, moved in a corresponding direction on the screen. Anticipation of these events may have affected neuronal activity during the previous part of the trial by inducing attention to the visual field periphery.

### 4.2.3 *Lever Task*

This task was used in Experiment 1. It was designed for a purpose outside the scope of the present study and consequently incorporates features that are irrelevant in the present context. However, one of its features – onset of one image against the backdrop of another already present image – allowed us to discover the phenomena at the center of this study. Accordingly, the description of the task will focus on the sequence in which images were presented. Events during representative trials for each of three conditions are shown in Figure 30A-C. While the monkey maintained central fixation, waiting for a target to appear, various task-irrelevant visual stimuli were presented, all centered on the fovea. Possible stimuli included the  $0.6^\circ$  blue fixation spot, a  $3.3^\circ \times 3.3^\circ$  gray square, a flanker array consisting of two  $1.5^\circ \times 0.6^\circ$  blue rectangles with their centers at an eccentricity of  $3.3^\circ$ , and the object, an image selected to elicit excitatory responses from the recorded neuron. The flankers never overlapped the object. In three conditions, the object and the flankers appeared in different temporal relations. In the 'object-alone' condition, the object appeared in the absence of the flanker (Figure 30A). In the 'flanker-then-object' condition, the object appeared against the backdrop of the already present flanker (Figure 30B). In the 'object-then-flanker' condition, the flanker appeared against the backdrop of the already present object (Figure 30C). The three conditions were imposed in separate blocks, each containing 192 successful trials. At the end of each trial, regardless of the prior sequence of stimuli, a target appeared. This was a  $0.7^\circ$  red disk placed  $5^\circ$  from the center of the screen at one of four locations: up and to the right, up and to the left, down and to the right or down and to the left (Figure 31). The monkey had to respond by moving one of two levers affixed to the primate chair

forward or backward. Moving the right (or left) lever forward (or backward) resulted in the right (or left) flanker moving upward (or downward). If the monkey moved a flanker so that it hit the target (Figure 31), he received a drop of juice. This phase of the trial is relevant to the present study only insofar as it presumably induced the monkey to allocate attention to the flankers.

#### *4.2.4 Fixation Task*

This task was used in Experiment 2. It was designed to test the hypothesis that the phenomena observed in Experiment 1 depended on the sequence in which the object and flanker were presented independently of their task relevance. Accordingly, it required the monkey simply to maintain central fixation during the presentation of task-irrelevant visual stimuli. Successful maintenance of fixation was rewarded with a drop of juice at the end of each trial. Four conditions were imposed in pseudorandom interleaved sequence until 16 trials had been completed successfully under each condition. In the "object-alone" condition, the object was displayed in isolation for 600 ms (Figure 30D). In the "flanker-then-object" condition, the flanker was visible for 600 ms and then the object was displayed against its backdrop for an additional 600 ms (Figure 30E). In the "flanker-alone" condition, the flanker was displayed in isolation for 600 ms (Figure 30F). In the "object-then-flanker" condition, the object was visible for 600 ms and then the flanker was displayed against its backdrop for an additional 600 ms (Figure 30G). The object was selected from the same library of 107 stimuli as in Experiment 1. The flanker was a red annulus with an inner radius of  $7.1^\circ$  and an outer radius of  $7.5^\circ$ . The centers of both coincided with the center of the screen.

#### *4.2.5 Statistical Analysis*

The next three sections describe the statistical procedures carried out on the data collected during the performance of the two tasks described above. First we assessed the visual responsiveness of neurons (Assessment of Visual Responsiveness). Neurons that were determined to be visually responsive were then included in two subsequent analyses. The first of these involves fitting a model to the time-course of the neuronal response (Model-Based Analysis). The parameters of this model were then used to assess the oscillations in terms of frequency, initial amplitude and damping time-constant. The second analysis was used to complement and confirm the results of the first. It involved using Fourier analysis to determine the power of the various frequencies in the neural signal (Fourier Analysis).

#### Assessment of Visual Responsiveness

To determine whether a neuron was visually responsive, we compared firing rates before and after onset of the object in the object-alone condition (paired  $t$ -test with criterion of  $p < 0.05$ ). In Experiment 1, the pre-stimulus epoch occupied the 200 ms immediately before object onset, while the post-stimulus epoch extended from 100 to 300 ms after object onset. In Experiment 2, the pre-stimulus epoch occupied the 550 ms immediately before object onset, while the post-stimulus epoch extended from 50 to 600 ms after object onset.

### Model-Based Analysis

For each visually responsive neuron under each experimental condition, we assessed the oscillatory component of the response by means of a curve-fitting procedure carried out with a commercially available package (Origin, MicroCal Software, Inc.). The period under consideration, as defined relative to stimulus onset, was 100-710 ms for the lever task and 100-610 ms for the fixation task. A function was fit to points representing mean firing rate vs. time in 10 ms bins. The function was  $F(t) = F_a(t) + F_b(t) * F_c(t)$ , where  $t$  represents time, with  $t = 0$  at 100 ms post-stimulus-onset, and where the individual terms are:

$$\text{Adapting Non-Oscillatory Component:} \quad F_a(t) = K_1 + \{K_2 / \exp[(t - K_3) / K_4]\}$$

$$\text{Gain and Adaptation of Oscillatory Component:} \quad F_b(t) = K_5 / \exp[(t - K_6) / K_7]$$

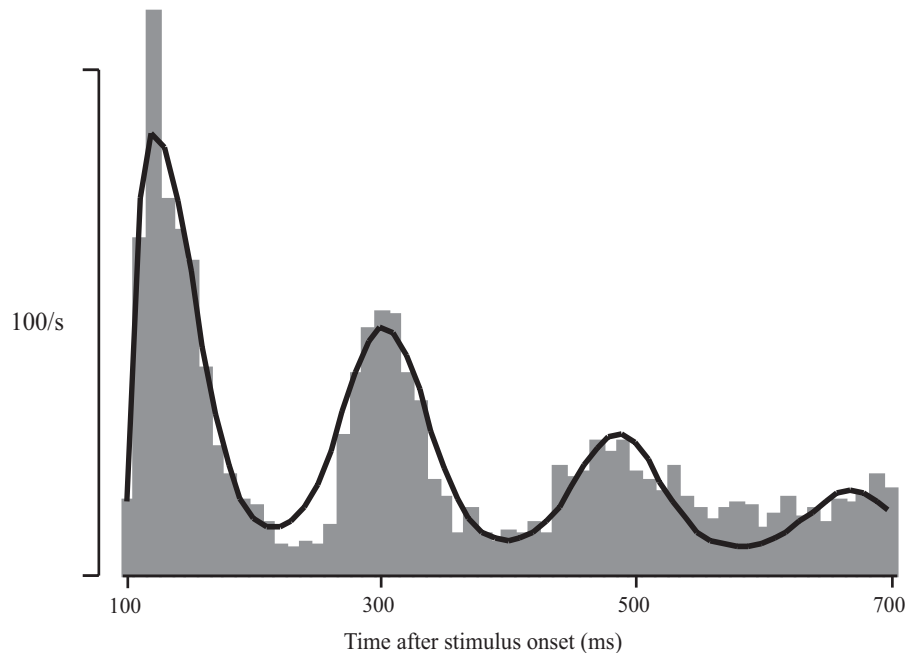
$$\text{Oscillatory Component:} \quad F_c(t) = \{2 + \cos[(t - K_8) / K_9]\}^{K_{10}}$$

Beginning the period of observation at a time close to the peak of the visual response (100 ms following stimulus onset) obviated having to model the onset of the response. Term  $F_a(t)$  captured the tendency for the response to wane exponentially even in the absence of any oscillatory component. Term  $F_b(t)$  allowed for independent adaptation of the oscillatory component. In term  $F_c(t)$ , adding 2 to the cosine function constrained the oscillation to vary between 1 and 3. Taking the resulting value to a variable power allowed controlling the sharpness of the peaks as compared to the troughs. This function gave a good fit to obviously oscillatory responses (Figure 32).

*Correlation Between Oscillatory Term and Histogram.* For each visually responsive neuron under each experimental condition, the curve-fitting procedure yielded a solution including parameters  $K_8$ ,  $K_9$  and  $K_{10}$  of  $F_c(t)$ , the oscillatory term. We



classified a response as potentially oscillatory if the frequency of the oscillatory term of the best-fit equation was greater than 4 Hz. We imposed this conservative criterion so as to rule out cases in which the fit might have arisen from an approximate match (within the 610 or 510 ms measurement epoch) between the oscillatory function and a phasic but non-periodic peak or trough in the response. To rate the degree of fit to the data afforded by the best-fit oscillatory term, we computed, for each neuron under each condition, the coefficient of correlation between this term and the histogram. We began with raw 10-ms-binned measures of mean firing rate:  $Y(0)$ ,  $Y(10)$ , ...,  $Y(T)$ , where  $Y(t)$  represents the mean firing rate in the interval between times  $t$  and  $t+10$  and where  $T = 600$  and  $500$  for Experiment 1 and 2 respectively. Then we conditioned the binned measures by removing variance that could be accounted for by the non-oscillatory  $F_a(t)$  and  $F_b(t)$  terms of the best-fit function. To do this, we used the formula:  $Y'(t) = [Y(t) - F_a(t)] / F_b(t)$ , with



**Figure 32. Example of curve-fitting from model-based analysis.** Post-stimulus-time histogram from Figure 33B together with curve fit to the data according to a procedure described in Methods. In Experiment 1, as shown here, curves were fit to data from a 610 ms period beginning 100 ms after stimulus onset. In Experiment 2, the duration of the sampling period was 510 ms.

parameters  $K_1$ - $K_7$  set to the values that had yielded an optimal fit. Then we measured the correlation across  $t$  between  $Y'(t)$  and the oscillatory function  $F_c(t)$ , with parameters  $K_8$ - $K_{10}$  of this function set to the values that had yielded an optimal fit. The resulting correlation coefficient was used for two purposes. First, it provided an indication of the goodness with which the oscillatory term fitted the data. Second, it served as a criterion for including or rejecting data in subsequent steps of analysis aimed at parametric characterization of oscillatory activity. If, for a given neuron under a given experimental condition, the correlation coefficient was greater than 0.25 (in Experiment 1 with 61 binned firing rate measures) or 0.27 (in Experiment 2 with 51 binned firing rate measures), then the response was classified as oscillatory and the data were included in subsequent stages of analysis. The threshold values were those which would, in a standard correlation analysis based on the respective numbers of observations, have yielded a significance level of  $p < 0.05$ . This significance level should not be taken literally because the observed and fitted measures were not independent. It was simply used as a means for equating the goodness-of-fit thresholds used in the two experiments.

*Parametric Characterization of Oscillatory Activity.* For each visually responsive neuron under each experimental condition in which it met the criterion for oscillatory activity (see previous section), we derived from the values of the best-fit parameters three attributes of the oscillatory best-fit function: the frequency, the initial amplitude and the damping time constant. These were computed from the parameters  $K_5$ - $K_{10}$  according to the following formulas. The frequency (Hz) was given by  $1000 / [2\pi * \text{abs}(K_9)]$ . The initial peak-to-peak amplitude (spikes/s) at  $t = 0$  was given by  $F_b(0) * (3^{K_{10}} - 1)$ . The damping time constant (ms) was given by  $K_7$ .

*Comparison of Distributions of Parameters.* To determine whether a measure of oscillatory activity (frequency, amplitude or damping time constant) differed significantly between experimental conditions, we carried out the following steps of analysis. First, we computed the distribution of values obtained in all visually responsive neurons that met the criterion for oscillatory activity under a given condition. Then we compared the distributions obtained under planned pairs of conditions. If the two distributions were not significantly different from normal (KS-Lilliefors test) then they were compared by a  $t$ -test. Otherwise, they were compared by a Mann-Whitney U-test. A layered Bonferroni correction took into account the occurrence of multiple pair-wise comparisons. This was not a within-neuron analysis. A given neuron might meet the criterion for oscillatory activity under one experimental condition and fail to meet it under another with the result that between-condition comparisons involved non-identical sets of neurons.

#### Fourier Analysis

In the model-based approach described above, ten parameters were adjusted so as to produce an optimal fit between the output of the model and the visual response histogram. Many of these parameters concerned properties of the visual response other than those of central concern (the strength and frequency of the oscillatory component of the response). Adjunctive properties placed under parametric control included the strength and time constant with which the firing rate was damped, the strength and time constant with which oscillatory activity was damped and the phase of the oscillatory activity. The advantage of being able to estimate these response properties was offset by a possible disadvantage. Because equal weight was given to each of the ten parameters,

the properties of central concern may not have been estimated as accurately as possible. To compensate for this limitation, we adopted an alternative approach focused exclusively on the strength and frequency of oscillatory activity. In this approach, we first constructed an autocorrelogram based on the interspike intervals and then, by Fourier analysis, computed the autocorrelogram's power spectrum. Comparing power spectra obtained under different experimental conditions allowed determining whether oscillatory activity at particular frequencies varied across conditions. An additional advantage of this method is that it considers each spike train separately, and thus avoids the assumptions present in the model-based analysis that oscillations are time locked to visual response onset and that visual response onset varies little from trial to trial.

Construction of an autocorrelogram representing the activity of a given neuron under a given experimental condition proceeded according to the following steps. Consideration was restricted to a limited period 100-700 ms after stimulus onset (in Experiment 1) or 100-600 ms after stimulus onset (in Experiment 2). Proceeding one trial at a time, and considering every pairwise combination of spikes within the measurement period, we accumulated counts of interspike intervals in 1 ms bins ranging from the minimal measurable interval (1 ms) to the maximal measurable interval (600 ms in Experiment 1 and 500 ms in Experiment 2). Even in the absence of low-frequency oscillations, an autocorrelogram constructed by measuring spikes in a finite temporal window will not be flat. For example, in the case of a neuron firing regularly at 1000 Hz, measuring interspike intervals in a 500 ms measurement period will result in 500 counts in the 1 ms bin, one count in the 500 ms bin, and a linear decline from 500 to 1 in the counts assigned to intervening bins. To compensate for this effect, we normalized the

count in each bin to the value  $(w - i + 1 \text{ ms})$ , where  $w$  was the duration of the measurement window and  $i$  was the duration of the interspike interval associated with the bin. Following this normalization, any deviation from flatness must reflect either noise or the presence of low frequency oscillatory activity. The depth of modulation of an autocorrelogram constructed according to the above steps depends on an accidental factor – the number of spikes in the full measurement period – determined by the mean firing rate and the number of trials. To convert from a count measure (dependent on these factors) to a frequency measure (independent of them), we divided the value of each bin by the number of spikes summed across the measurement windows of all trials.

Prior to computation of the power spectrum, we conditioned the histogram so as to minimize the impact of the discontinuities at its edges. First, we rendered it symmetrical by combining the histogram with its mirror image formed by reflection across the 0 ms bin. Then we multiplied it by a Gaussian of the form  $y = \exp(-i^2/2\sigma^2)/(2\pi\sigma)^{1/2}$  where  $i$  = bin number (-500 to +500) and  $\sigma = 250$  (Experiment 1) or 200 (Experiment 2). Finally, we zero-padded it to a factor of  $2^{13}$  and carried out a Fourier transformation.

In characterizing the power spectrum obtained for a given neuron under a given condition, we first determined whether power attained a genuine maximum (with first derivative of zero) in the 4-7 Hz range. If a maximum could be identified, we then determined at what frequency it occurred and recorded its power. Subsequent analyses were based on these values. For example, in order to determine whether oscillatory strength differed significantly between two conditions, we carried out a paired  $t$ -test on

the measured maximum of power in the 4-7 Hz range for all neurons in which a genuine maximum was present under both conditions.

## 4.3 Results

### 4.3.1 Overview

This study began with an incidental observation concerning oscillatory neuronal visual responses in IT of monkey Op. The monkey had learned to perform a task in which, during steady fixation, an object effective at eliciting neuronal activity was presented at the fovea (Figure 30A). In some trials, flanking bars were already present on the screen at the time when the object appeared (Figure 30B). We noted on several occasions that the neuronal visual response, as rendered on the audio monitor, developed an obvious 'chatter' whenever the flankers were visible. We set out, in Experiment 1, to document and extend this observation by recording from IT neurons in monkey Op under three separate conditions presented in blocks. The purpose of the object-alone condition was to characterize the baseline visual response to the object (Figure 30A). The purpose of the flanker-then-object condition was to characterize any increase in oscillatory activity induced when the object was presented against the backdrop of the flanker (Figure 30B). A third condition, object-then-flanker (Figure 30C), was included to explore the possibility, suggested by certain models (see Discussion), that phase-reversed oscillatory responses might occur when the flanker was presented against the backdrop of the object. We collected data from 65 neurons in monkey Op under all three conditions. The results confirmed our initial observation that presenting an object against the backdrop of a flanker induced or enhanced an initially positive oscillatory response.

Furthermore, they confirmed our speculation that presenting the flanker against the backdrop of the object would induce an initially negative oscillatory response.

The task used in Experiment 1 probably induced the monkey to divert attention from the object, which was task-irrelevant, to the visual field periphery, where the target was going to appear and where the flankers, which served as proxies for the manipulanda, were sometimes visible (Figure 31). To determine whether the dependence of oscillatory activity on a background image was specific to this behavioral context or could be observed when there was no pressure to allocate attention to the periphery, we carried out Experiment 2. We trained monkeys Op and Fi simply to maintain central fixation while the object and the flanker were presented in various sequences. A new flanker (an annulus centered on fixation) was used so as to prevent any association with the first task. A new condition (flanker alone) was included as a baseline for the analysis of oscillatory activity elicited when the flanker was presented against the backdrop of the object. The square which had appeared before the object in Experiment 1 (Figure 30A, panel A2) was eliminated because it was irrelevant to the analysis of oscillatory activity. The sequence of events occurring under each of four interleaved conditions of Experiment 2 is portrayed in Figure 30D-G. In the context of this task, we recorded from 75 neurons in monkey Op and 103 neurons in monkey Fi. A small set of data from a third monkey was collected during the fixation task and included in the fourier analysis in experiment 2 (monkey Ph:  $n = 21$ ).

#### 4.3.2 Location of Recording Sites

Recording was carried out in anterior IT in the right hemisphere of two monkeys (monkeys Op and Fi) and the left hemisphere of a third monkey (monkey Ph). In all three animals, recording sites were lateral to the anterior medial temporal sulcus. Recording in monkey Op was confined to frontal levels in the range anterior 18-22 mm as defined with respect to the interaural plane. The range of recording sites was 17-20 mm in monkey Fi, and 13-16 mm in monkey Ph. With respect to depth, recording sites in monkey Op were limited to the ventral aspect of the inferotemporal gyrus, whereas recording sites in monkeys Fi and Ph were localized to the lower bank of the superior temporal sulcus as well as the ventral aspect of the inferotemporal gyrus. There was no obvious trend toward variation in neuronal properties with respect to the location of the recording site.

#### 4.3.3 Experiment 1: Example of Oscillatory Visual Response

Data collected from one neuron recorded from during the lever task are shown in Figure 33A-C. For this cell, the presence of the flanker clearly enhanced the oscillatory component of the visual response to the preferred object. When the object was presented alone, this neuron gave a response in which there was a slight oscillatory tendency as

---

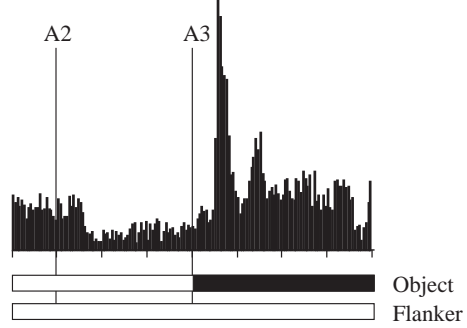
**Figure 33. Example neurons from Experiments 1 and 2.** A-C. Activity of a single neuron under the three conditions imposed in Experiment 1. Each histogram represents average firing rate as a function of time for 192 trials corresponding to the condition indicated. Event markers (A2, A3, B2, B3, C2 and C3) refer to the onset of displays with corresponding labels in Figure 30A-C. D-G. Activity of a single neuron under the four conditions imposed in Experiment 2. Each histogram represents the average firing rate as a function of time for 16 trials corresponding to the condition indicated. Event markers (D2, D3, E2, E3, F2, F3, G2 and G3) refer to the onset of displays with corresponding labels in Figure 30D-G.

---

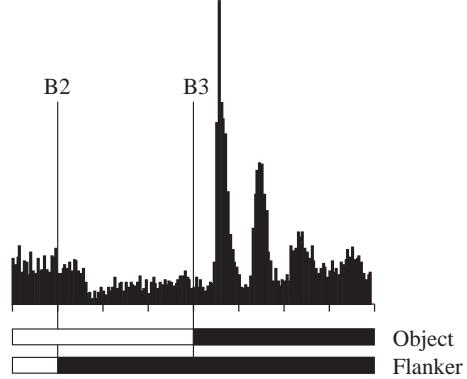


## Experiment 1

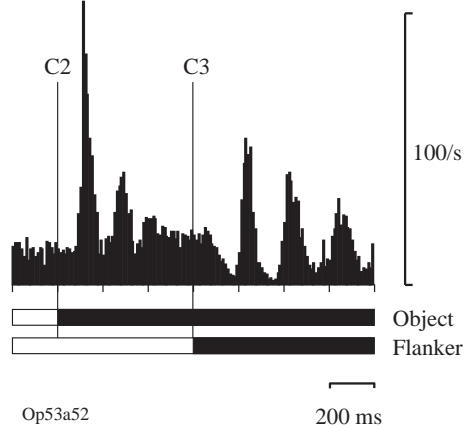
### A. Object Alone



### B. Flanker then Object

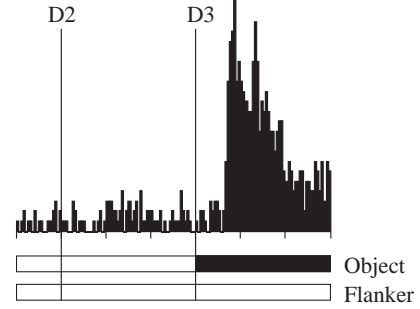


### C. Object then Flanker

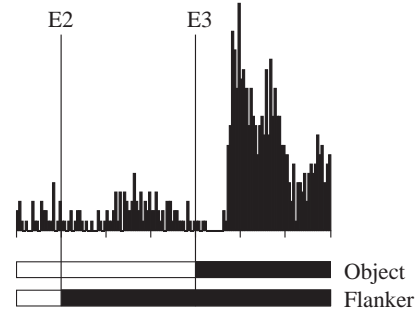


## Experiment 2

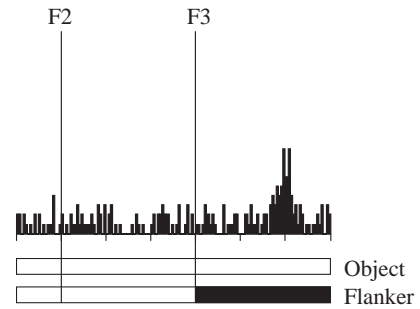
### D. Object Alone



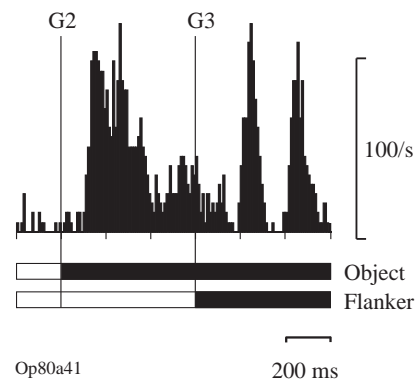
### E. Flanker then Object



### F. Flanker Alone



### G. Object then Flanker



indicated by the occurrence of a second peak around 200 ms after the first peak (Figure 33A). When the object was presented against the backdrop of an already present flanker, the tendency of the neuron to oscillate was markedly enhanced. This was evident by an increase in the peak-to-trough amplitude and a prolongation such that up to four peaks were discernible (Figure 33B). When the flanker appeared against the backdrop of the already present object, a dramatic oscillatory response also occurred, but in this case took the form of an initial phase of suppressed activity followed by a series of peaks and troughs (Figure 33C). The robust oscillations occurring under this condition are all the more striking by contrast to the complete lack of any oscillatory tendency in the response to onset of the flanker and square in the absence of the object (Figure 33B, event B2).

**Table 4. Numbers of neurons included in the model-based analysis.** Condition: object-alone (O), flanker-then-object (F-O), flanker-alone (F), and object-then-flanker (O-F). Rows labeled 1-4 contain counts of neurons meeting progressively more stringent criteria. 1) Recorded: neurons from which a full set of data was collected under all conditions in a given task. 2) Visual: neurons satisfying the above condition and giving a significant visual response to the object in the object-alone condition. 3) freq > 4 Hz: neurons satisfying the above conditions and in which the frequency of the oscillatory term of the best-fit function was greater than 4 Hz. 4)  $r > .25$  or  $.27$ : neurons satisfying the above conditions and in which the coefficient of correlation between the best-fit oscillatory term and the histogram was greater than  $.25$  (in Experiment 1) or  $.27$  (in Experiment 2). These coefficients represent the degree of correlation expected by change at a probability of  $.05$  in light of the numbers of bins in the histograms (61 in Experiment 1 and 51 in Experiment 2). Analysis under “Rate of Incidence of Oscillatory Activity” (see text) was based on neurons in row 3. Analysis under “Parameters of Oscillatory Activity” (see text) was based on neurons in row 4.

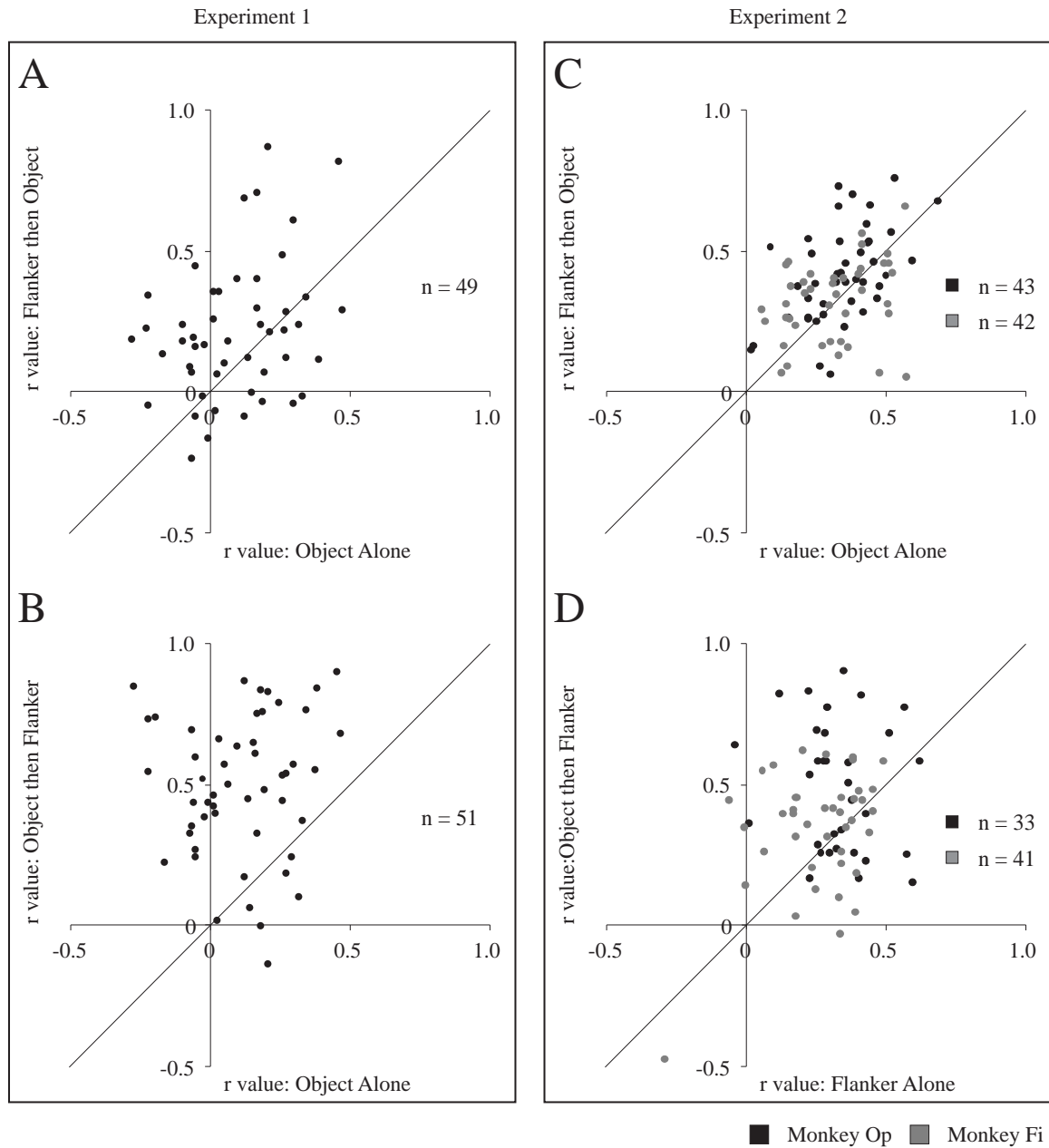
Monkey	Monkey Op							Monkey Fi			
Task	Experiment 1 Lever Task			Experiment 2 Fixation Task				Experiment 2 Fixation Task			
Condition	O	F-O	O-F	O	F-O	F	O-F	O	F-O	F	O-F
1) Recorded	65			75				103			
2) Visual	62			58				71			
3) freq > 4 Hz	54	53	59	46	49	34	57	52	55	54	56
4) r > .25 or r > .27	13	18	47	33	39	25	49	33	35	29	37

#### 4.3.4 Experiment 1: Model-Based Analysis

##### Rate of Incidence of Oscillatory Activity

Of the 65 neurons from which data were collected (Table 4, row 1), 62 gave statistically significant visual responses to the object in the object-alone condition (Table 4, row 2). Only these 62 neurons were considered further. We assessed the oscillatory activity of each neuron under each of the three experimental conditions by means of a curve-fitting procedure (see Methods and Figure 32). Thus curves were fit to 186 histograms (3 conditions x 62 neurons). We classified a response as potentially oscillatory if the frequency of the oscillatory term of the best-fit function was greater than 4 Hz (see Methods for rationale). This criterion was met by 54, 53 and 59 cases out of 62 under the object-alone, flanker-then-object and object-then-flanker conditions respectively (Table 4, row 3). Only these cases were considered further. In each case, we estimated the goodness of fit between the histogram and the oscillatory term of the best-fit function by computing a correlation coefficient (see Methods).

To determine whether neurons exhibited more pronounced oscillations in response to stimuli turned on in the presence of other stimuli than in response to stimuli turned on in isolation, we compared correlation coefficients between appropriate pairs of conditions. Flanker-then-object vs. object-alone. Forty-nine neurons met the criterion that the frequency of the oscillatory term be greater than 4 Hz under both of the conditions to be compared. We carried out a within-neuron comparison of the correlation coefficients in these cases. The results are presented in Figure 34A, in which each point represents a neuron and in which the point's location with respect to the horizontal (or vertical) axis represents the correlation coefficient obtained under the object-alone (or



**Figure 34. Measures of oscillatory activity.** Within-neuron comparison of measures of oscillatory activity obtained during presentation of stimuli under different conditions in Experiment 1 (A-B) and Experiment 2 (C-D). Each point represents a neuron. The position of each point with respect to a given axis is determined by the correlation coefficient reflecting the goodness of fit of the oscillatory term of the best-fit function to the histogram for the corresponding condition. A. Flanker-then-object vs. object-alone. B. Object-then-flanker vs. object-alone. C. Flanker-then-object vs. object-alone. D. Object-then-flanker vs. flanker-alone. The oscillatory index tended to be greater under conditions in which a stimulus was presented against a backdrop, with the result that the majority of points fell above the identity line. This effect achieved significance in all cases except that of monkey 2 in graph C (see text).

flanker-then-object) condition. That some correlation coefficients were negative reflects the curve-fitting procedure's having stopped short of an optimal solution in cases where a large amount of the variance in the histogram could be accounted for by the non-oscillatory terms. There was a highly significant tendency for the correlation coefficient to be greater under the flanker-then-object condition, as reflected by the preponderance of points above the identity line (paired  $t$ -test,  $p < 0.005$ ). Object-then-flanker vs. object-alone. Fifty-one neurons met the criterion that the frequency of the oscillatory term be greater than 4 Hz under both of the conditions to be compared. For these neurons, the scatter plot of Figure 34B compares correlation coefficients obtained under the object-alone condition (horizontal axis) to correlation coefficients obtained under the object-then-flanker condition (vertical axis). There was a strong and highly significant tendency for the correlation coefficient obtained under the object-then-flanker condition to be greater (paired  $t$ -test,  $p < 0.0001$ ). We conclude that oscillatory activity was more pronounced when either the object or the flanker was presented against the already visible backdrop of the other image than when the object was presented alone.

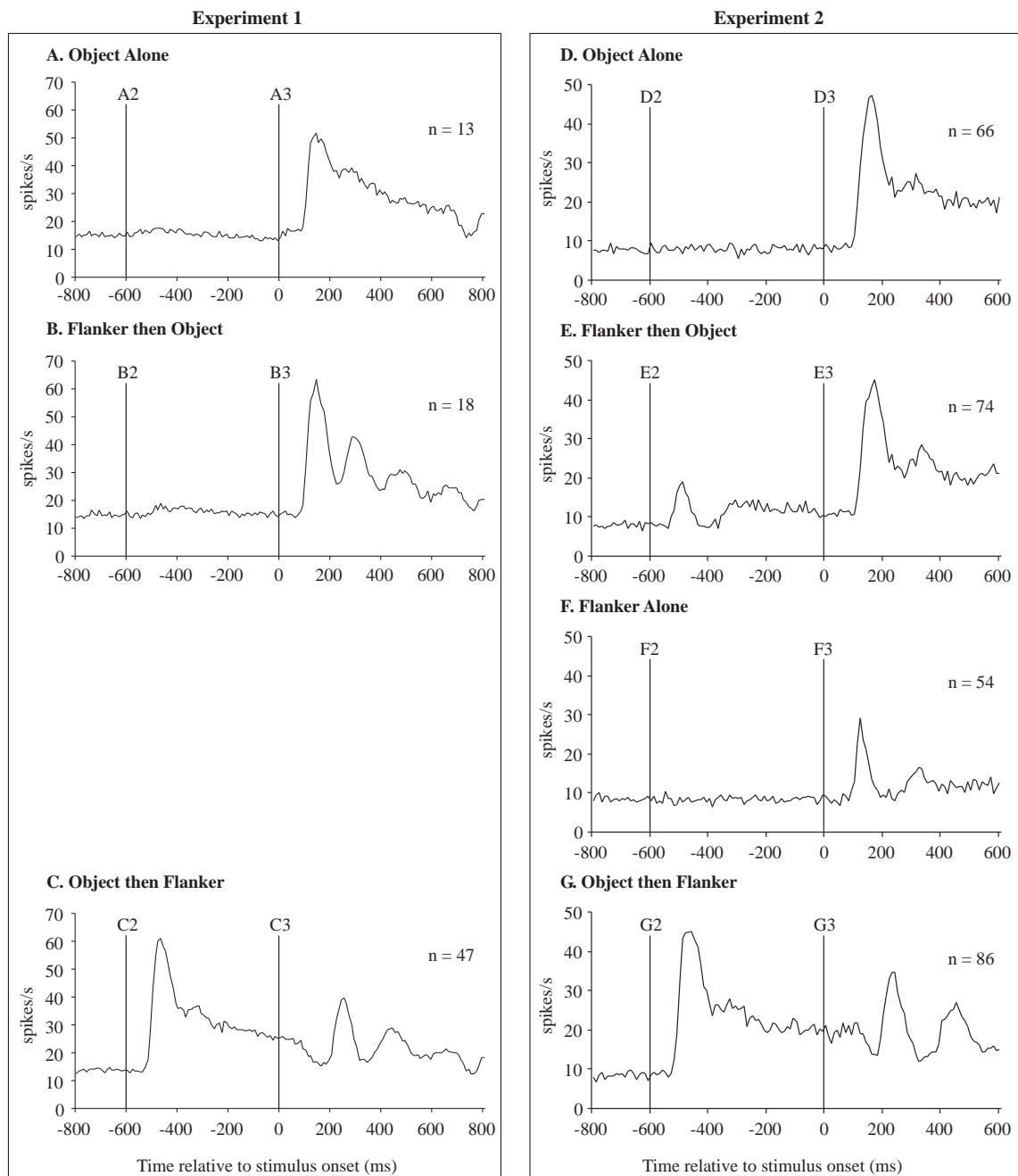
#### Parameters of Oscillatory Activity

We next sought to determine, in cases in which oscillatory activity occurred, whether its properties were different in the presence of a backdrop (flanker-then-object and object-then-flanker conditions) than in the absence of a backdrop (object-alone condition). To eliminate cases in which the occurrence of oscillatory activity was questionable, we considered only those cases in which the correlation coefficient was greater than 0.25 (see Methods for rationale). This criterion was met in 13, 18 and 47

neurons in the object-alone, flanker-then-object and object-then-flanker conditions respectively (Table 4, row 4). With the data from these neurons, we then created population histograms, and calculated the frequency, amplitude, and damping time constant of the oscillatory activity.

*Population Histograms.* For all neurons meeting the criteria for oscillatory activity described above, we created population histograms representing the average firing rate as a function of post-stimulus time. These histograms are shown in Figure 35A-C. It is important to note that these histograms do not represent the responses of the entire sampled population of neurons, but rather represent the responses of subpopulations of neurons exhibiting oscillatory activity under particular conditions. Having constructed the histograms, we could then ask whether oscillatory activity, when present, was qualitatively different across conditions. Under the object-alone condition, oscillatory activity was restricted to a subtle decrease in firing rate following the initial phasic component of the visual response (Figure 35A). In contrast, oscillatory activity was pronounced under the flanker-then-object (Figure 35B) and object-then-flanker (Figure 35C) conditions. Between these conditions, however, the oscillatory activity differed in phase, beginning with a peak under the flanker-then-object condition (Figure 35B) and with a trough in the object-then-flanker condition (Figure 35C).

To test the impression that oscillations were greater in amplitude and more prolonged when a stimulus was presented against a visible backdrop, we carried out a series of quantitative analyses, using the parameters of the curves fit to the individual-case histograms to estimate the frequency, amplitude and damping time constant of oscillatory activity. For each parameter, we compared (1) the distribution in object-alone



**Figure 35. Population histograms.** Histograms represent mean firing rate as a function of time for all neurons meeting the criteria for oscillatory visual responsiveness under a given condition (Table 4, row 4). A-C. Three conditions from Experiment 1. D-G. Four conditions from Experiment 2. The alignment event, indicated by the vertical line at time 0, was onset of the object in A-B and D-E. It was onset of the flanker in C and F-G. Event markers (A2, A3... G2, G3) refer to the onset of displays thus labeled in Figure 30.

to that in flanker-then-object and (2) the distribution in object-alone to that in object-then-flanker.

*Frequency.* We first computed the frequency of the oscillatory term of the best-fit function (see Methods). The resulting distributions of frequencies are shown in Figure 36A-C. The means of the distributions for the object-alone, flanker-then-object and object-then-flanker conditions were 5.8, 5.8, and 5.2 Hz, respectively. Comparison of the distributions in object-alone to those in flanker-then-object and object-then-flanker revealed no significant difference. We conclude that if oscillatory activity occurred then its frequency was in a range centered between 5 and 6 Hz regardless of the presence or absence of a visible backdrop at the time of stimulus presentation.

*Amplitude.* The mean (across all cases in each experimental condition) of oscillatory amplitude as measured at the beginning of the response (see Methods) is shown in Figure 37A. The means for flanker-then-object and object-alone were significantly different (Mann-Whitney U-test,  $p < 0.02$ ), with the mean amplitude greater for flanker-then-object. Thus, even with consideration restricted to cases that met the criterion for oscillatory activity, the oscillatory component was stronger when the object was presented against the backdrop of the flanker than when the object was presented alone.

---

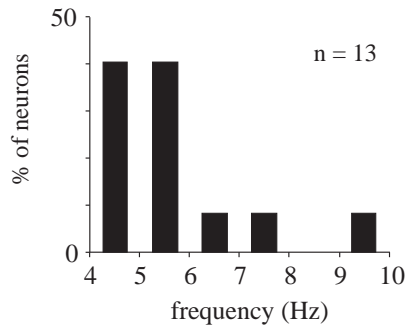
**Figure 36. Frequency of oscillations.** Distributions of measured frequency of oscillatory activity for all neurons meeting the criteria for oscillatory activity under each condition in each experiment (Table 4, row 4). A-C. Three conditions in Experiment 1. D-G. Four conditions in Experiment 2. In neither experiment was the distribution of frequencies different when a given stimulus was presented against a backdrop than when it was presented alone (see text).

---

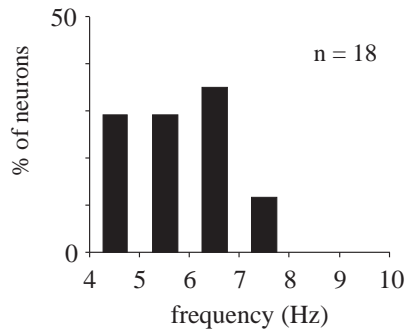


## Experiment 1

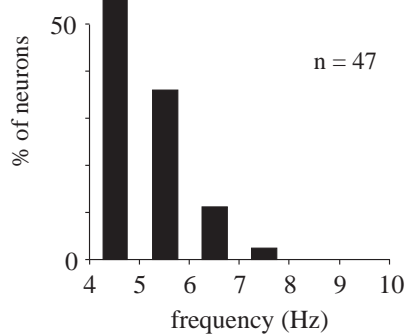
### A. Object Alone



### B. Flanker then Object

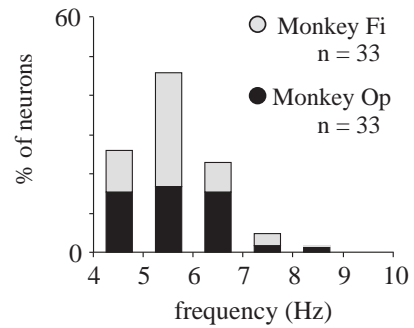


### C. Object then Flanker

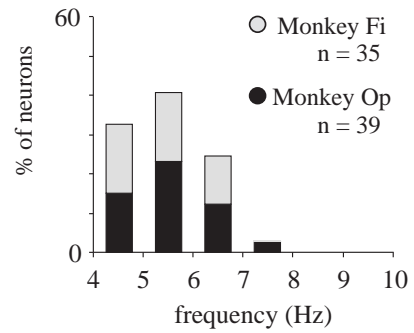


## Experiment 2

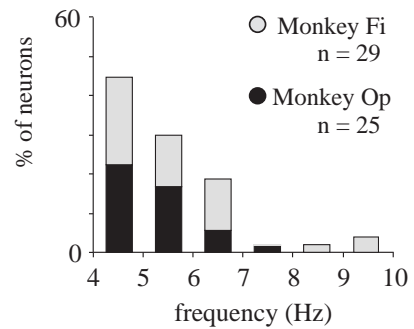
### D. Object Alone



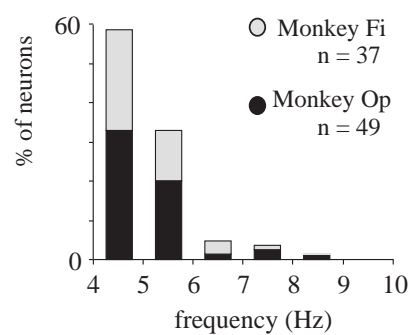
### E. Flanker then Object



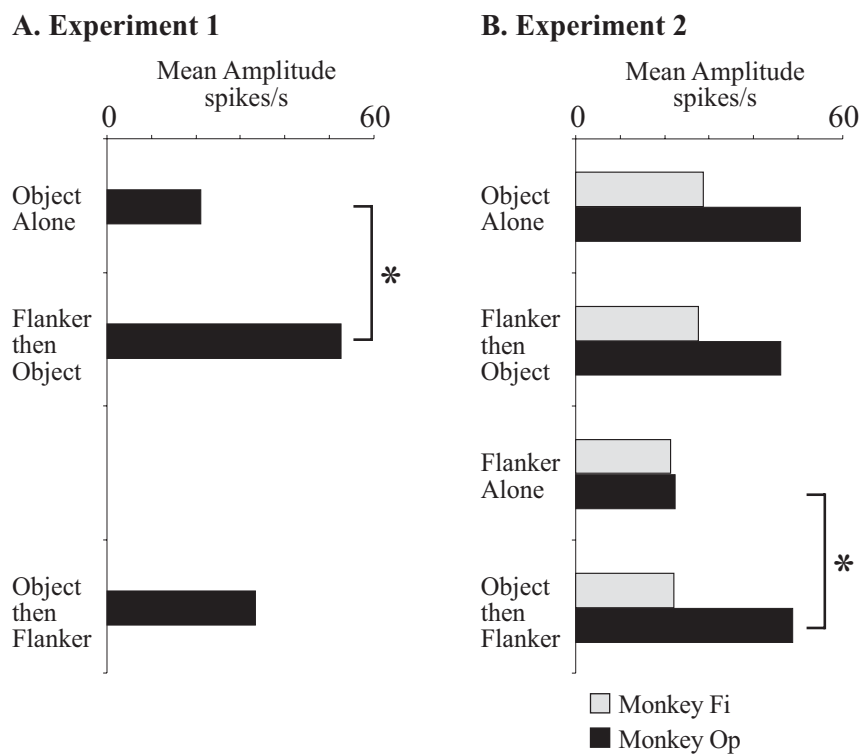
### F. Flanker Alone



### G. Object then Flanker



*Damping Time Constant.* The distributions of the damping time constants (see Methods) are shown in Figure 38A-C. The means of the distributions for the object-alone, flanker-then-object and object-then-flanker conditions were 507, 325, and 1,818 ms respectively. The difference between the object-alone and flanker-then-object conditions was not significant. However, the difference between the object-alone and object-then-flanker conditions did attain significance (Mann-Whitney U-test,  $p < 0.02$ ). Thus, even with consideration restricted to cases that met the criterion for oscillatory activity, the oscillatory component died down more slowly when the flanker was presented against the backdrop of the object than when the object was presented alone.



**Figure 37. Mean amplitude of oscillatory activity.** Amplitude was measured at the outset of the response, prior to decay, for all neurons meeting the criteria for oscillatory activity under each condition in each experiment (Table 4, row 4). A. Three conditions in Experiment 1. B. Four conditions in Experiment 2. In two comparisons involving monkey 1 (asterisks), the mean amplitude was significantly greater when a stimulus was presented against a backdrop than when it was presented in isolation (see text).

#### 4.3.5 Experiment 1: Fourier Analysis

##### Example of Oscillatory Visual Response

Auto-correlograms (ACGs) and power spectra of data collected from the example neuron shown in Figure 33A-C are presented in Figure 39A-D. The slight tendency for this neuron to oscillate in response to the object presented alone is reflected in a peak in the ACG at approximately  $\pm 150$  ms (Figure 39A). As demonstrated with the raw histogram data, the ACGs show that the presence of the flanker clearly enhanced the oscillatory component of the visual response to the preferred object. The increase in the peak to trough amplitude as well as the prolongation of oscillations in the flanker-then-object condition (Figure 33B) is reflected in the enhancement of and the increase in number of peaks in the ACG (Figure 39B). When the flanker appeared against the backdrop of the already present object, a dramatic oscillatory response also occurred, but at a slightly lower frequency than the other two conditions (Figure 39C). To quantify these observations, the data in the ACGs were Fourier transformed, with the resulting power spectra presented in Figure 39D. Red, green and blue lines represent the power spectra for the object-alone, flanker-then-object and object-then-flanker conditions, respectively. In all three conditions, the oscillations reflected in the ACGs appear as distinct peaks in the power spectra at approximately 5 Hz. There exists greater power at

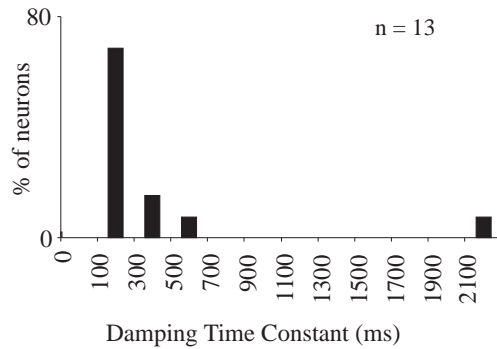
---

**Figure 38. Time-constants of oscillatory decay.** Distributions of estimated time-constants for all neurons meeting the criteria for oscillatory activity under each condition in each experiment (Table 4, row 4). A-C. Three conditions in Experiment 1. D-G. Four conditions in Experiment 2. The distribution in C was shifted significantly to the right relative to that in A; that in G was shifted significantly to the right relative to that in F (see text). These effects reflect slower damping of oscillatory activity when a stimulus was presented against a backdrop than when it was presented in isolation.

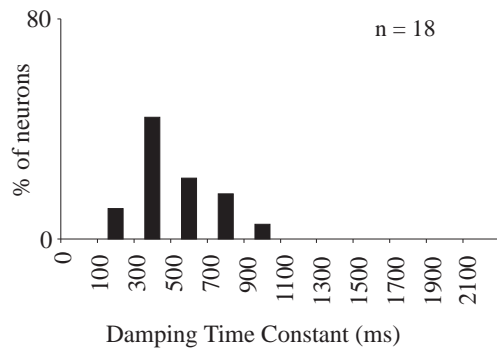
---

## Experiment 1

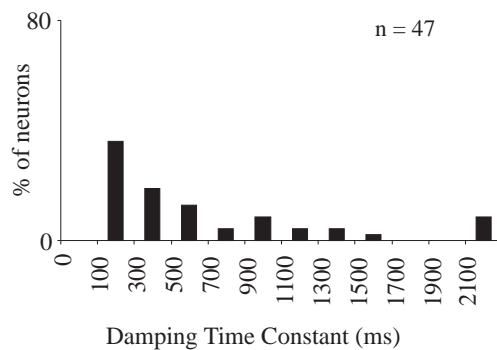
### A. Object Alone



### B. Flanker then Object

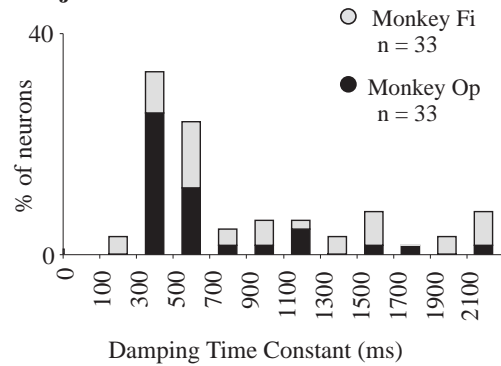


### C. Object then Flanker

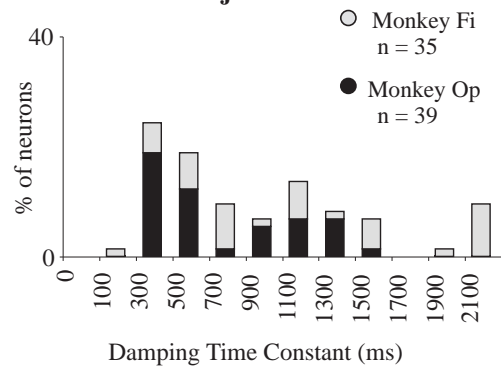


## Experiment 2

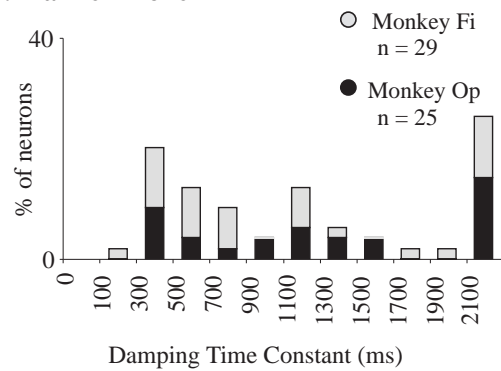
### D. Object Alone



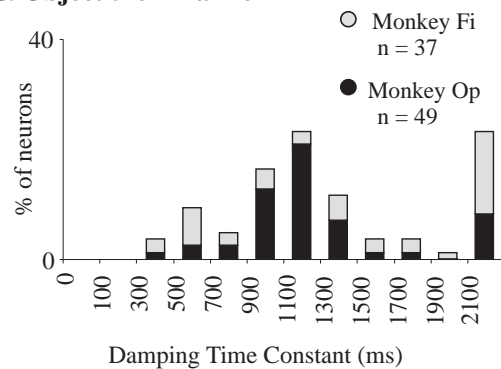
### E. Flanker then Object



### F. Flanker Alone



### G. Object then Flanker



this frequency when stimuli were turned on in the presence of other stimuli than when stimuli were turned on in isolation as demonstrated by the higher spectral peaks for the flanker-then-object and object-then-flanker conditions. Moreover, as observed in the ACGs, the frequency of oscillations was slightly lower for the object-then-flanker condition than the other two conditions.

### Population Power Spectra

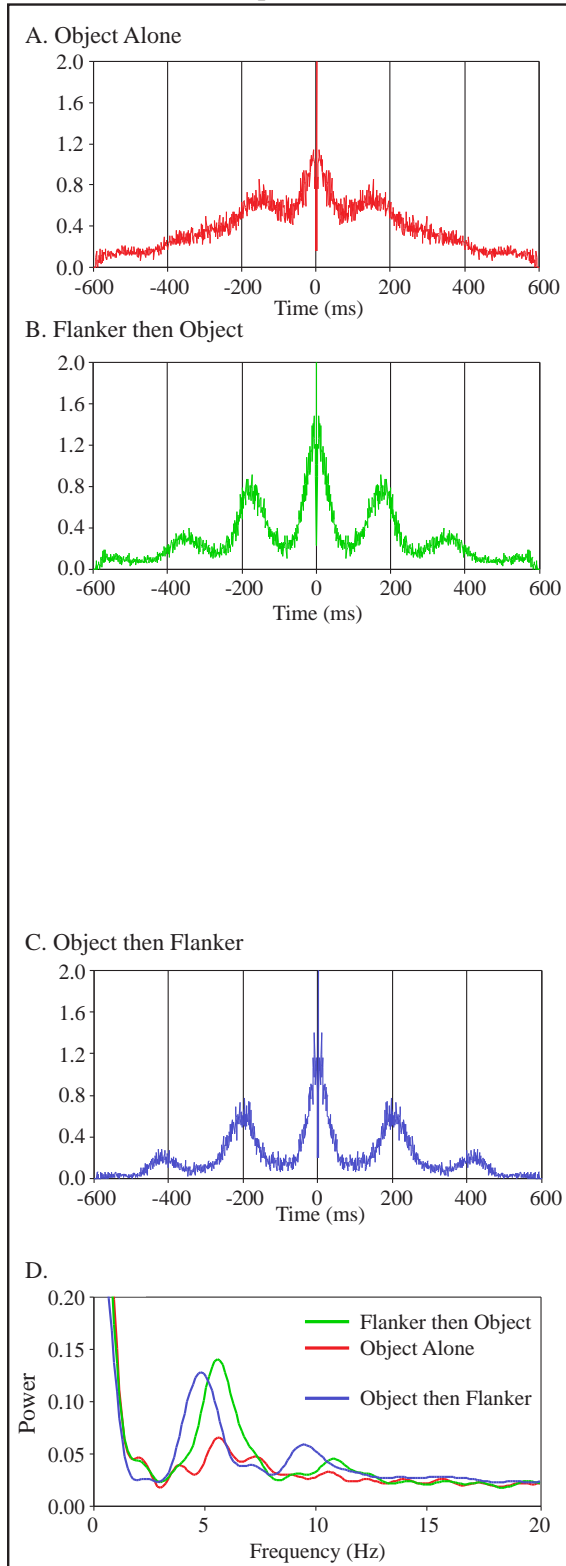
For all 62 visually responsive neurons, we created a population power spectrum representing the average power as a function of frequency for each of the three conditions. These power spectra are shown in Figure 40A. Red, green and blue lines represent average power spectra for object-alone, flanker-then-object and object-then-flanker conditions, respectively. Since there were no obvious differences in the power spectra for the three conditions for frequencies greater than 15 Hz, the power spectra for frequencies only up to 15 Hz are shown. The power for low frequency oscillations in the range of 5-6 Hz was greater in the flanker-then-object and object-then-flanker conditions than the object alone condition. These results show that, as observed in the example neuron and the model-based analysis of the population of neurons, the oscillations were

---

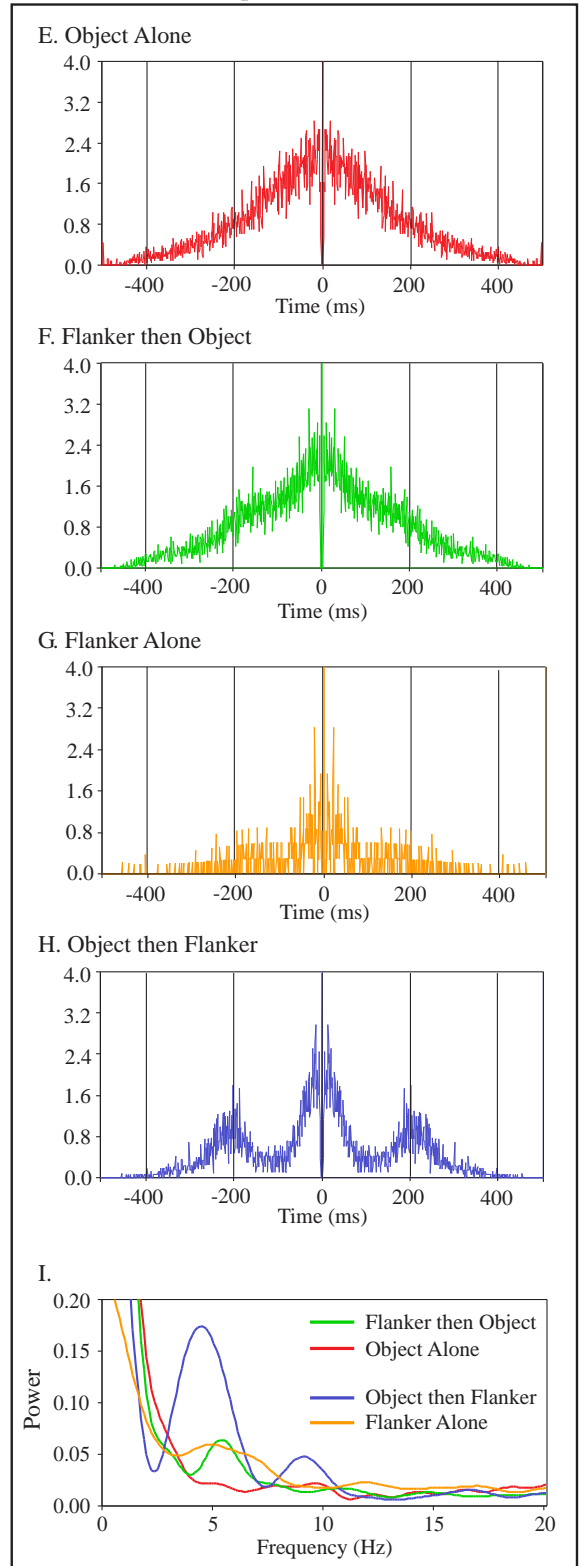
**Figure 39. Auto-correlograms (ACGs) and power spectra of data from the two example neurons shown in Figure 33.** A-C. ACGs for the three conditions imposed in Experiment 1. Plots were calculated over a 600 ms epoch, multiplied with a Gaussian and normalized to the number of spikes summed across the measurement windows in all trials. D. Power spectra of data from three conditions in Experiment 1. E-H. ACGs for the four conditions imposed in Experiment 2. Plots were calculated over a 500 ms epoch. I. Power spectra of data from four conditions in Experiment 2. Red, green, orange and blue lines represent data from the object-alone, flanker-then-object, flanker-alone and object-then-flanker conditions, respectively.

---

## Experiment 1



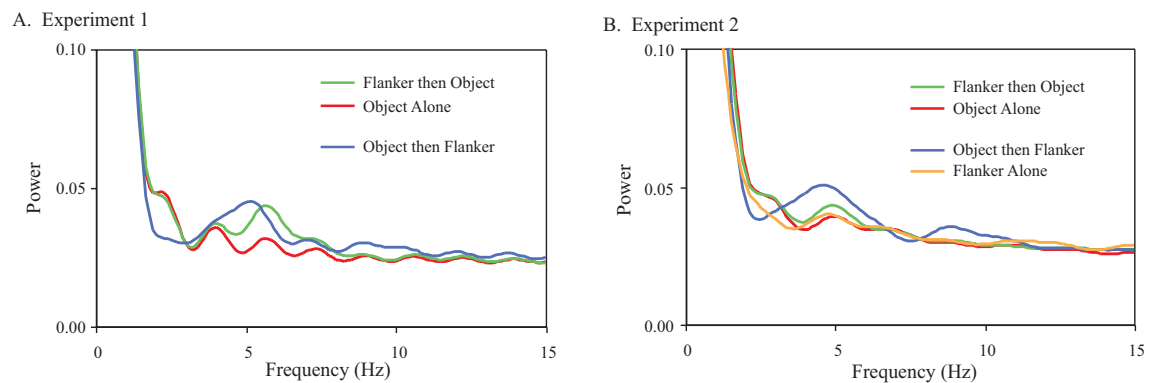
## Experiment 2



more pronounced when stimuli were turned on in the presence of other stimuli than when stimuli were turned on in isolation.

### Within Neuron Comparisons of Oscillatory Responses

To assess the statistical significance of these trends, we compared the power of low frequency oscillations in the power spectra between appropriate pairs of conditions. Flanker-then-object vs. object-alone. For 57 neurons, the power spectra contained a peak between 4 and 7 Hz under both the Flanker-then-object and object-alone conditions. We carried out a within-neuron comparison of the values of power in these cases. The results are presented in Figure 41A, in which each point represents a neuron and in which the point's location with respect to the horizontal (or vertical) axis represents the peak spectral power in the 4-7 Hz range for the object-alone (or flanker-then-object) condition. The spectral power was significantly greater for the flanker-then-object condition, as



**Figure 40. Population power spectra.** Power spectra were averaged across 62 neurons in Experiment 1 (A) and 149 neurons in Experiment 2 (B). There was greater power at low frequencies when stimuli were turned on in the presence of other stimuli (green and blue lines) than when stimuli were turned on in isolation (red and orange lines). Red, green, orange and blue lines represent power spectra of data collected in the object-alone, flanker-then-object, flanker-alone and object-then-flanker conditions, respectively.

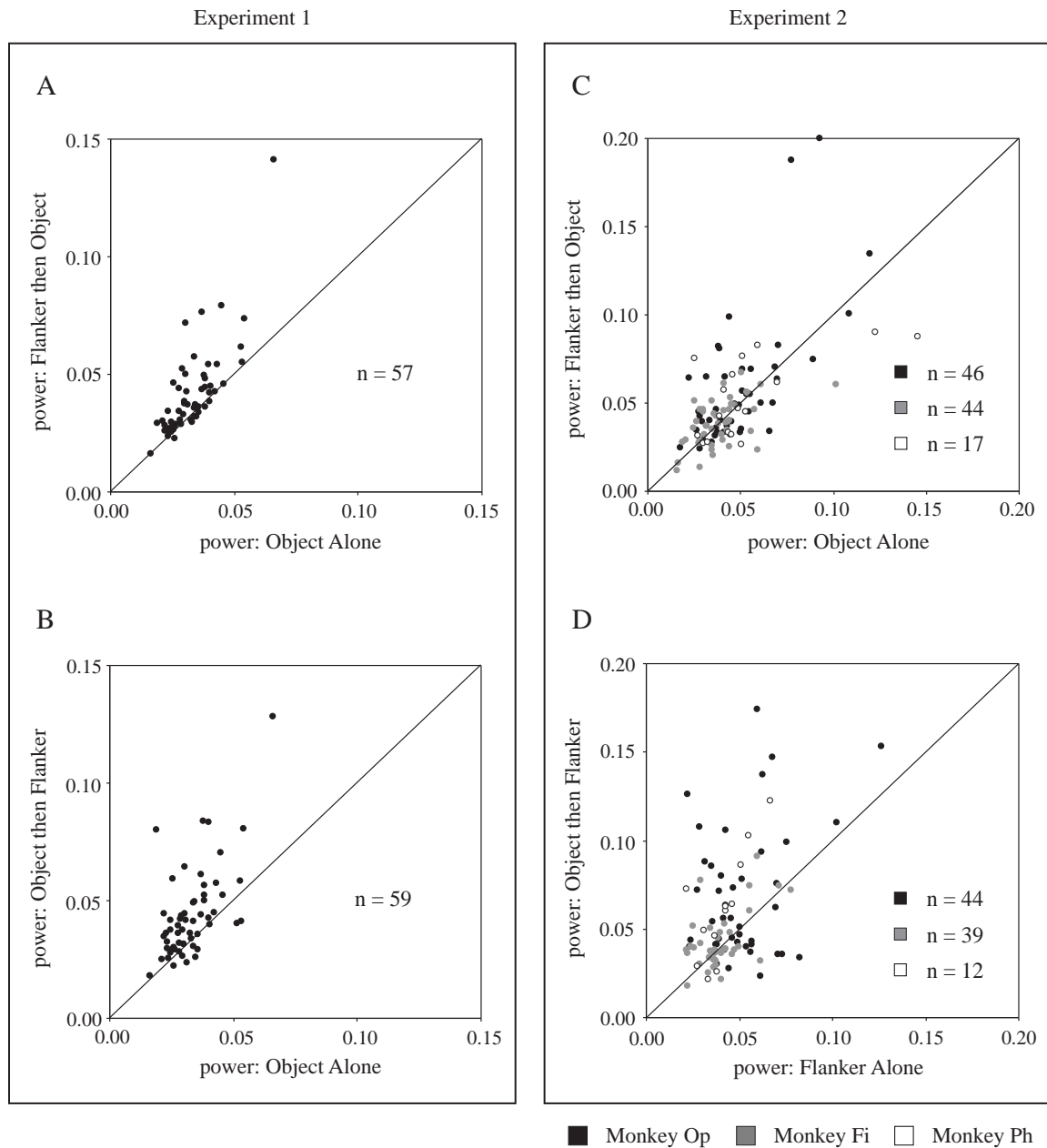
reflected by the preponderance of points above the identity line (paired  $t$ -test,  $p < 0.000001$ ). Object-then-flanker vs. object-alone. For 59 neurons, the power spectrum

contained a peak between 4 and 7 Hz under both the object-then-flanker and object-alone conditions. For these neurons, the scatter plot of Figure 41B compares the peak spectral power in the 4-7 Hz range for the object-alone condition (horizontal axis) to the peak spectral power in the 4-7 Hz range for the object-then-flanker condition (vertical axis). The spectral power was significantly greater for the object-then-flanker condition, as reflected by the preponderance of points above the identity line (paired  $t$ -test,  $p < 0.000001$ ). We conclude, in confirmation of results obtained with the curve fitting analysis, that oscillatory activity was more pronounced when an image was presented against a backdrop than when the image was presented alone.

#### *4.3.6 Experiment 2: Example of Oscillatory Visual Response*

The activity of one neuron studied in Experiment 2 is presented in Figure 33D-G. Data from this neuron demonstrated that oscillatory visual responses occurred even without the necessity of making a lever response to a peripheral target. While presentation of the object in isolation elicited a strong response in which there was little or no oscillatory tendency (Figure 33D), when the object was presented against the backdrop of the flanker, a moderately oscillatory pattern emerged (Figure 33E). Presentation of the flanker in isolation elicited no response at the standard latency (Figure 33F). However, presentation of the flanker against the backdrop of the object elicited a dramatic initially negative oscillatory response (Figure 33G).





**Figure 41. Within neuron comparisons of spectral power in the frequency range of 4-7 Hz.** Comparisons were made for conditions in Experiment 1 (A-B) and Experiment 2 (C-D). Each point represents a neuron. The position of each point with respect to a given axis is the power of the highest peak in the power spectra within the frequency range of 4-7 Hz for the corresponding condition. A. Flanker-then-object vs. object-alone. B. Object-then-flanker vs. object-alone. C. Flanker-then-object vs. object-alone. D. Object-then-flanker vs. flanker-alone. The power at low frequencies tended to be greater under conditions in which a stimulus was presented against a backdrop. This effect achieved significance for both comparisons in Experiment 1 and for monkey Op and Ph in graph D.

#### 4.3.7 Experiment 2: Model-Based Analysis

##### Rate of Incidence of Oscillatory Activity

In monkeys Op and Fi, respectively, 58 and 71 neurons gave statistically significant visual responses to the object in the object-alone condition (Table 4, row 2). We measured the oscillatory activity of each of these neurons under each of the four stimulus conditions by the same curve-fitting procedure as was used in Experiment 1 (see Methods). We considered further only cases in which the frequency of the oscillatory term of the best-fit function was greater than 4 Hz (see Methods for rationale). In monkey Op, 46, 49, 34 and 57 cases met this criterion under the object-alone, flanker-then-object, flanker-alone and object-then-flanker conditions respectively; the corresponding counts in monkey Fi were 52, 55, 54 and 56 (Table 4, row 3). For each of these cases, we estimated the goodness of fit by measuring the coefficient of correlation between histogram and the oscillatory term of the best-fit function (see Methods).

To determine whether neurons exhibited more pronounced oscillations when a stimulus was turned on in the presence of another stimulus than when it was turned on in isolation, we compared the values of the correlation coefficients between selected pairs of experimental conditions. Flanker-then-object vs. object-alone. Forty-three neurons in monkey Op and 42 neurons in monkey Fi met the criterion that the frequency of the best-fit oscillatory term be greater than 4 Hz under both of the conditions to be compared. We carried out a within-neuron comparison of the correlation coefficients in these cases. The results are presented in Figure 34C, in which each point represents a neuron and in which the point's location with respect to the horizontal (or vertical) axis represents the correlation coefficient obtained under the object-alone (or flanker-then-object) condition.

The correlation coefficient tended to be greater under the flanker-then-object condition, as reflected by the preponderance points above the identity line. This tendency was significant in monkey Op (paired  $t$ -test,  $p < 0.01$ ) but not in monkey Fi. Object-then-flanker vs. flanker-alone. Thirty-three neurons in monkey Op and 41 neurons in monkey Fi met the criterion that the frequency of the best-fit oscillatory term be greater than 4 Hz under both of the conditions to be compared. For these neurons, the scatter plot of Figure 34D compares correlation coefficients obtained under the flanker-alone condition (horizontal axis) to correlation coefficients obtained under the object-then-flanker condition (vertical axis). There was a marked tendency for the correlation coefficient obtained under the object-then-flanker condition to be greater than under the flanker-alone condition. This effect was significant in both monkeys (paired  $t$ -test; monkey Op:  $p < 0.01$ ; monkey Fi:  $p < 0.05$ ). We conclude, in confirmation of results obtained in Experiment 1, that oscillatory activity was more pronounced when an image was presented against a visible backdrop than when it was presented alone.

#### Parameters of Oscillatory Activity

The aim of the next step of the analysis was to determine, in cases in which oscillatory activity occurred, whether its properties were different in the presence versus in the absence of a backdrop (flanker-then-object vs. object-alone condition; object-then-flanker vs. flanker-alone condition). To eliminate cases in which the occurrence of oscillatory activity was questionable, we considered only those in which the correlation coefficient was greater than 0.27 (see Methods for rationale). In monkey Op, this criterion was met in 33, 39, 25 and 49 neurons in the object-alone, flanker-then-object,

flanker-alone and object-then-flanker conditions respectively; in monkey Fi, the corresponding counts were 33, 35, 29 and 37 (Table 4, row 4). On the basis of these cases, population histograms were constructed and measurements of frequency, amplitude and the damping time constant were carried out.

*Population Histograms.* We constructed population histograms representing average firing rate as a function of post-stimulus time for all neurons meeting the criteria for oscillatory activity described in the previous paragraph. These histograms, shown in Figure 35D-G, allowed us to ask whether oscillatory activity, when present, was qualitatively different across conditions. Under the object-alone condition, oscillatory activity of the population was restricted to a subtle dip in firing rate following the initial phasic component of the visual response (Figure 35D). This pattern was slightly more pronounced under the flanker-then-object condition (Figure 35E). Under the flanker-alone condition, there was a weak phasic excitatory response followed by a dip and rebound (Figure 35F). Finally, under the object-then-flanker condition, oscillatory activity was pronounced and began with a trough rather than a peak (Figure 35G). To characterize oscillatory activity further, we carried out a series of quantitative analyses, using the parameters of the curves fit to the individual-case histograms to estimate the frequency, amplitude and damping time constant of oscillatory activity. For each parameter and in each monkey, we compared (1) the distribution in object-alone to that in flanker-then-object and (2) the distribution in flanker-alone to that in object-then-flanker.

*Frequency.* The distributions of frequencies are shown in Figure 36D-G. The means of the distributions were between 5 and 6 Hz in both monkeys under all conditions. In monkey Op, the means were 5.7, 5.5, 5.2 and 5.2 Hz under the object-

alone, flanker-then-object, flanker-alone and object-then-flanker conditions respectively. In monkey Fi, the corresponding values were 5.6, 5.3, 5.7 and 5.1 Hz. Tests comparing the distributions obtained under flanker-then-object vs. object-alone condition and object-then-flanker vs. flanker-alone condition revealed no significant differences. We conclude, in confirmation of results obtained in Experiment 1, that if oscillatory activity occurred, then its frequency was in a range centered between 5 and 6 Hz regardless of the presence or absence of a visible backdrop at the time of stimulus presentation.

*Amplitude.* The mean (across all cases in each experimental condition) of oscillatory amplitude as measured at the beginning of the response (see Methods) is shown in Figure 37B. In monkey Op, the mean amplitude was significantly greater under the object-then-flanker condition than under the flanker-alone condition (Mann-Whitney U-test,  $p < 0.0002$ ). No other comparison yielded a significant result. Thus, even with consideration restricted to cases that met the criterion for oscillatory activity, the oscillatory component was stronger when the flanker was presented against the backdrop of the object than when it was presented alone. This replicates for flanker the enhancement of amplitude resulting from presentation of the object against a backdrop in Experiment 1 (Figure 37A).

*Damping Time Constant.* The distributions of damping time constants are shown in Figure 38D-G. One comparison yielded a significant result: in monkey Fi, the distribution of time constants was shifted toward higher values under the object-then-flanker as compared to the flanker-alone condition (Mann-Whitney U-test,  $p < 0.05$ ). Thus, even with consideration restricted to cases that met the criterion for oscillatory activity, the oscillatory component died down more slowly when the flanker was

presented against the backdrop of the object than when it was presented alone. An analogous effect was observed in Experiment 1 on comparing object-then-flanker to object-alone.

#### *4.3.8 Experiment 2: Fourier Analysis*

##### Example of Oscillatory Visual Response

Auto-correlograms (ACGs) and power spectra of data collected from the example neuron shown in Figure 33D-G are presented in Figure 39E-I. This neuron did not show any tendency to oscillate in response to the object presented alone, as reflected in the absence of peaks in the ACG in Figure 39E. The modest increase in oscillatory activity in the flanker-then-object condition compared to the object-alone condition is reflected in a slight bump in the ACG shown in Figure 39F at around  $\pm 150$  ms. The noise in the ACG for the flanker-alone condition (Figure 39G) reflects the low firing rate evoked by the flanker presented alone. In contrast to the other three conditions, the object-then-flanker condition showed marked oscillations, as reflected by the peaks at  $\pm 200$  ms in the ACG (Figure 39H).

The power spectra resulting from the Fourier transform of this data are presented in Figure 39I. Red, green, orange and blue lines represent the power spectra for the object-alone, flanker-then-object, flanker-alone and object-then-flanker conditions, respectively. It is clear that there is greater power for frequencies in the range of 5 Hz for the flanker-then-object condition compared to the object-alone condition. There is an even greater increase in power for frequencies around 5 Hz for the object-then-flanker

condition compared to the flanker-alone condition. These observations confirm those made for the ACG and curve fitting analyses.

### Population Power Spectra

For all 149 visually responsive neurons (monkey Op:  $n = 58$ ; monkey Fi:  $n = 71$ ; monkey Ph:  $n = 20$ ), we created a population power spectrum representing the average power as a function of frequency for each of the four conditions. These power spectra are shown in Figure 40B. Red, green, orange and blue lines represent average power spectra for object-alone, flanker-then-object, flanker-alone and object-then-flanker conditions, respectively. Since there were no obvious differences in the power spectra for the four conditions for frequencies greater than 15 Hz, the power spectra for frequencies only up to 15 Hz are shown. The power for low frequency oscillations in the range of 5-6 Hz was greater in the flanker-then-object condition than in the object-alone condition. There is an even greater increase in power for frequencies in the range of 5-6 Hz for the object-then-flanker conditions over the flanker-alone condition. These results show that for the fixation task, as observed in the example neuron and the curve fitting analysis of the population of neurons, the oscillations were more pronounced when one stimulus was turned on in the presence of another, than when a stimulus was presented in isolation.

### Within Neuron Comparisons of Oscillatory Responses

To assess the statistical significance of these trends, we compared the power of low frequency oscillations in the power spectra between appropriate pairs of conditions.

Flanker-then-object vs. object-alone. For 107 neurons, the power spectrum contained a peak between 4 and 7 Hz under both the flanker-then-object and object-alone conditions (monkey Op:  $n = 46$ ; monkey Fi:  $n = 44$ ; monkey Ph:  $n = 17$ ). We carried out a within-neuron comparison of the values of power in these cases for each monkey separately. The results are presented in Figure 41C, in which each point represents a neuron and in which the point's location with respect to the horizontal (or vertical) axis represents the peak spectral power in the 4-7 Hz range for the object-alone (or flanker-then-object) condition. Although there was a trend toward greater power at low frequencies for the flanker-then-object condition compared to the object-alone condition, this trend was not significant for any of the three monkeys (paired  $t$ -test; monkey Op:  $p > 0.1$ ; monkey Fi:  $p > 0.5$ ; monkey Ph:  $p > 0.6$ ). Object-then-flanker vs. flanker-alone. For 95 neurons, the power spectrum contained a peak between 4 and 7 Hz under both the object-then-flanker and object-alone conditions (monkey Op:  $n = 44$ ; monkey Fi:  $n = 39$ ; monkey Ph:  $n = 12$ ). For these neurons, the scatter plot of Figure 41D compares the peak spectral power in the 4-7 Hz range for the flanker-alone condition (horizontal axis) to the peak spectral power in the 4-7 Hz range for the object-then-flanker condition (vertical axis). The spectral power was significantly greater for the object-then-flanker condition for monkeys Op and Ph (paired  $t$ -test; monkey Op:  $p < 0.006$ ; monkey Fi:  $p > 0.3$ ; monkey Ph:  $p < 0.02$ ). We conclude, in confirmation of results obtained with the curve fitting analysis, that oscillatory activity was more pronounced when an image was presented against a backdrop than when the backdrop was presented alone.



#### 4.3.9 Additional Observations

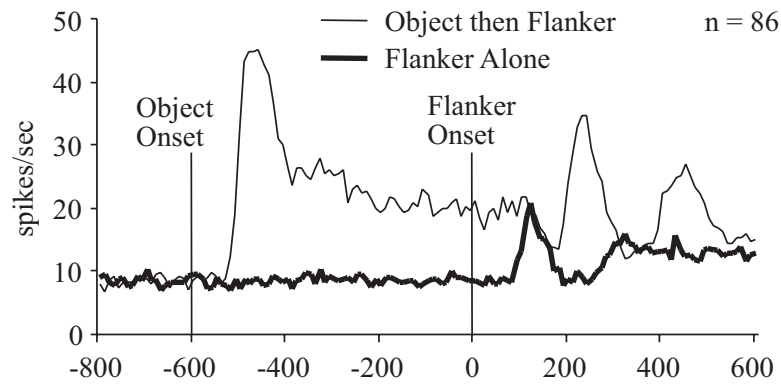
##### Experiment 1 vs. Experiment 2

The main aim of Experiment 2 was to determine whether phenomena observed in Experiment 1 would occur even when there was no pressure on the monkey to allocate attention to the peripheral visual field. We found, indeed, that the initially positive oscillatory response was enhanced under the flanker-then-object as compared to the object-alone condition (Figure 35D-E) and an initially negative oscillatory response emerged under the object-then-flanker as compared to the flanker-alone condition (Figure 35F-G). Nevertheless, there were hints of differences in outcome between the two experiments. In particular, in cases in which the flanker-then-object display elicited oscillatory activity, this activity appeared to be less pronounced and less prolonged in Experiment 2 (Figure 35E) than in Experiment 1 (Figure 35B). This relative diminution, although worth note, cannot be interpreted without further experiments. It might have arisen from the lack of necessity for peripheral attention, from other changes in task design, or from our having sampled a population of neurons with subtly different properties.

##### Oscillatory Frequency of Responses to Objects and Flankers

The distribution of frequencies appeared to be shifted toward lower values when onset of the flanker against the backdrop of the object elicited oscillatory activity (Figure 36C, G) than when onset of the object against the backdrop of the flanker did so (Figure 36B, E). This could also be seen in the population power spectra (Figure 40). *Post hoc* analysis revealed that the distributions (calculated with the curve fitting procedure) indeed differed significantly (Mann Whitney U; Experiment 1:  $p < 0.02$ ; Experiment 2:  $p$

$< 0.02$ ). The significance of this observation is not immediately clear. However, it provides a useful constraint on any effort to model oscillatory responses in detail.

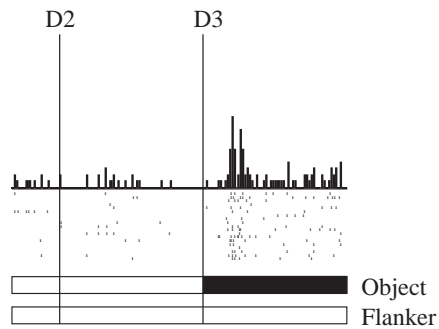


**Figure 42. Population histogram representing mean firing rate as a function of time under the flanker-alone condition and the object-then-flanker condition for all neurons exhibiting oscillatory visual responses under the latter condition.** This is the same population of 86 neurons on which Fig. 35G is based. Although the initial response to the flanker presented alone was excitatory, the initial response to the flanker presented against the backdrop of the object was suppressive.

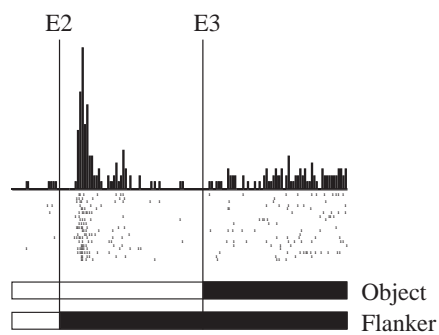
### The Effect of the Flanker: Competition vs. Inhibition

Oscillations elicited by presenting the flanker against the backdrop of the object began with a trough (Figure 35C, G). From this observation, one might be tempted to infer that the effect of the flanker was inhibitory and that the oscillations arose from alternating inhibition and excitatory rebound. However, this inference cannot be correct because the same population of neurons that responded with early suppression to the object-then-flanker display (Figure 42, thin curve) responded with early excitation to the flanker-alone display (Figure 42, thick curve). This observation suggests that the initially suppressive response in object-then-flanker arose from a form of competition in which the flanker, when presented against the backdrop of the object, drew population activity toward the lower (but still excitatory) firing rate elicited by the flanker in isolation. If this were so, then, in any neuron more responsive to the flanker than to the object, one

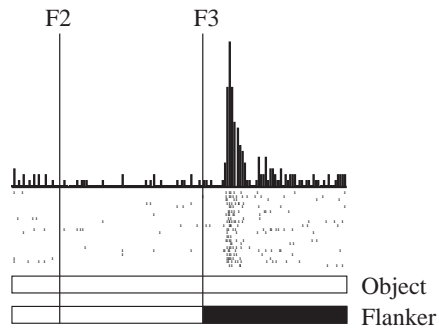
### A. Object Alone



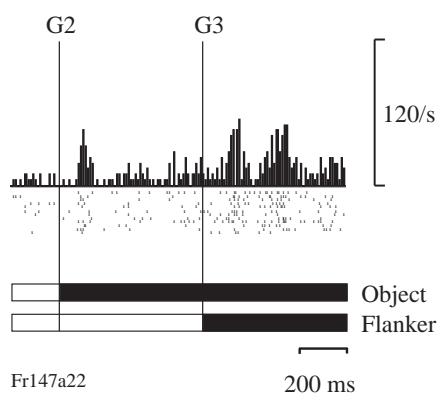
### B. Flanker then Object



### C. Flanker Alone



### D. Object then Flanker

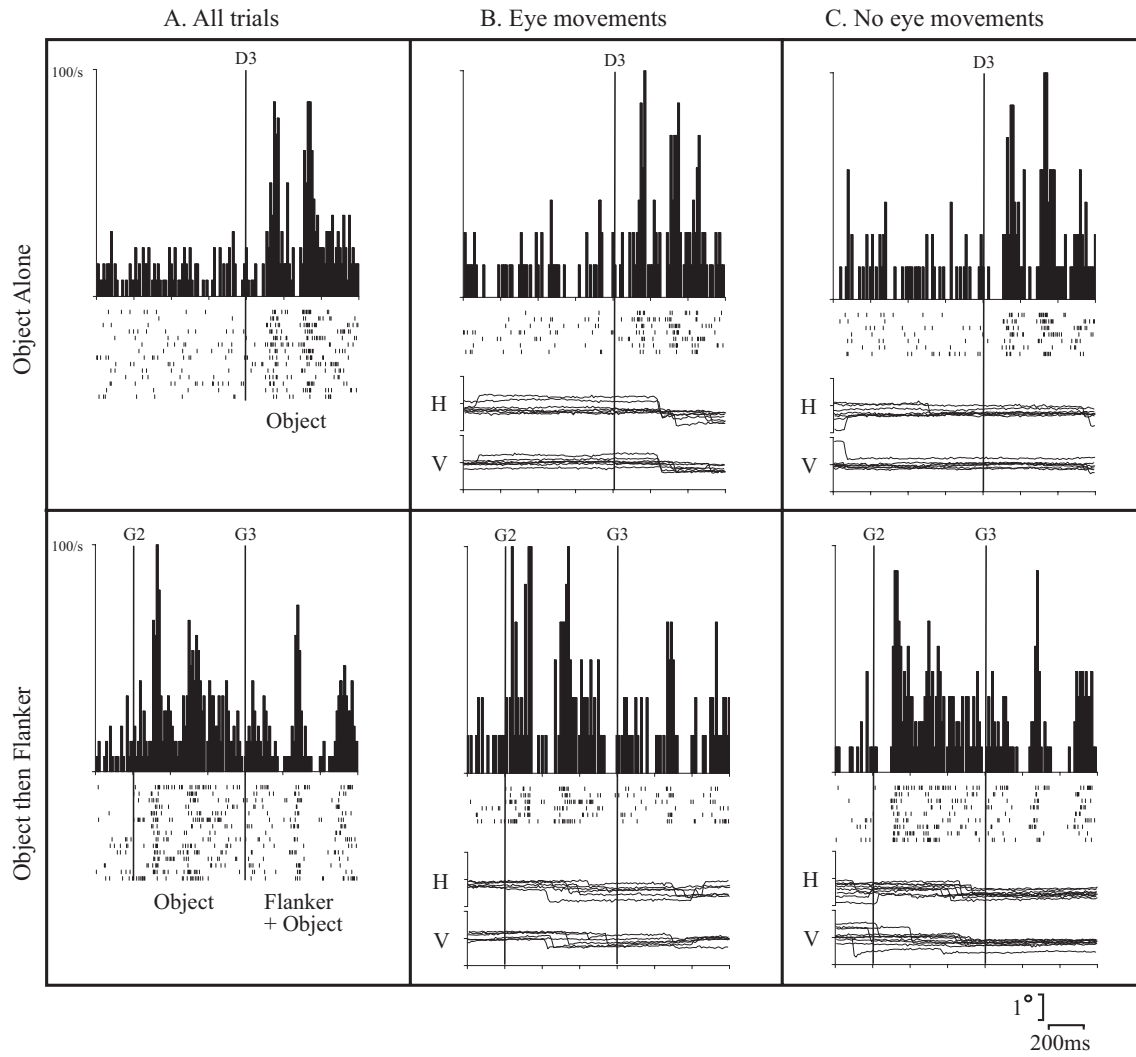


**Figure 43. Example of a neuron more responsive to flanker.** This neuron exhibited the unusual pattern of responding more strongly to the flanker (C) than to the object (A). When the flanker was displayed against the backdrop of the already present object, the response took the form of an initially positive oscillation (D) rather than the initially negative form observed in other neurons (Fig. 33G).

would expect that the oscillatory response under the object-then-flanker condition should have an initially positive phase. This is precisely what was observed in the neuron shown in Figure 43, which possessed the unusual property of responding better to the flanker than to the object. It thus appears that oscillations arose from competition between the flanker and the object to control the neuronal firing rate rather than from summation of an excitatory influence exerted by the object and an inhibitory influence exerted by the flanker.

### Eye Movements

Throughout all of the above experiments, eye position was continuously monitored and monkeys were required to maintain fixation within a window approximately  $1^\circ$  in diameter. Microsaccades were tolerated so long as they did not take the eye outside the window. To assess whether eye movements were correlated with oscillatory activity, we examined data from sessions in which oscillatory activity and eye position data had been stored. We found no relation between the two. Data supporting this point are presented in Figure 44. This neuron displayed initially positive oscillations under the object-alone condition and initially negative oscillations under the object-then-flanker condition (Figure 44A). Oscillatory activity was at least as strong and well defined when the period of oscillatory activity was devoid of microsaccades (Figure 44C) as on trials in which microsaccades occurred (Figure 44B).



**Figure 44. Data from a neuron giving an initially positive oscillatory response to the object presented in isolation (top row) and an initially negative oscillatory response to the flanker presented against the backdrop of the already present object (bottom row). A.** Oscillations are clearly visible both in the histogram representing mean firing rate as a function of time during the trial and in the raster displays representing activity during individual trials. **B.** Data from a subset of trials characterized by small deflections of the eye during the post-stimulus epoch. **C.** Data from the remaining trials, in which the eye was stable during the post-stimulus epoch. Note that oscillations are as prominent in the absence of eye movements (C) as in their presence (B). H and V represent horizontal and vertical position of the eye as a function of time during the trial. Eye traces from all trials in each set are superimposed.

## 4.4 Discussion

### 4.4.1 Overview

Some neurons in IT, presented with an effective foveal image, respond to it with a series of bursts at a frequency of approximately 5 Hz. The study reported here has yielded three novel observations related to this phenomenon. First, the strength of oscillatory activity is enhanced if the foveal stimulus is presented against the backdrop of an already present peripheral flanking display. Second, turning on the peripheral display against the backdrop of an already present foveal stimulus elicits an initially negative oscillatory response. Third, the occurrence of these phenomena is not critically dependent on task context: they occur both in a context promoting attention to the visual field periphery and in one requiring only central fixation.

### 4.4.2 Low-Frequency Oscillatory Activity in IT

Previous microelectrode recording studies of IT have revealed that some neurons respond to visual stimuli by firing rhythmically at a frequency of around 5 Hz. Nakamura et al. (1991, 1992), recording from the temporal pole, in subdivisions of TE and TG anterior to those studied here, documented oscillatory visual responses within this frequency range and observed a trend whereby familiar objects elicited stronger oscillations than unfamiliar ones (Nakamura et al., 1991). Cases of IT neurons with oscillatory activity in roughly this range have also been presented incidentally in Figure 5A of Sheinberg and Logothetis (1997), Figure 3E of Sato (1989), Figure 3A-B of Sato et al. (1980) and Figures 3-4 of Tamura and Tanaka (2001). Tovee and Rolls (1992), searching explicitly for oscillatory activity in IT, failed to obtain evidence for it even in

the low frequency range studied here. The reasons for this failure are not clear but may include their recording at sites roughly 1 cm posterior to those studied here.

#### *4.4.3 Stimulus-Stimulus Interactions in IT*

Several previous studies assessed the impact on neuronal activity in IT of presenting a neuron's preferred image simultaneously with another image, either a different image or a duplicate of the preferred image (Miller et al., 1993a; Missal et al., 1999; Rolls and Tovee, 1995; Sato, 1989, 1995). The essential finding is that supplementing an effective image with a second image leads to reduced responsiveness. In no case was oscillatory activity noted. In most cases the onset of the two images was simultaneous, rather than staggered as in the present study. However, in one case, even staggered presentation did not yield oscillatory activity (Sato, 1995). The discrepancy between that result and ours might be related to any of several factors. First, recording sites in the former study extended rostrally only to the middle of the anterior middle temporal sulcus, whereas recording sites in our study were located almost exclusively rostral to this level. Second, the stimulus set employed in the former study (consisting of seven geometric figures and four colored spots) was more limited than ours (consisting of 107 relatively complex images) with the result that we might have been able to match the stimulus preferences of recorded neurons more closely. Third, stimuli employed in the former study were both located in the peripheral visual field, whereas, in our study, a small preferred stimulus at the fovea was balanced against a large non-preferred stimulus at an eccentric location. It is not clear which of these factors is important. Outstanding questions relevant to this issue include the following: (1) do oscillations occur even at

relatively posterior levels in IT; (2) do oscillations arise from competition between neurons representing different images or representing different locations and, (3) insofar as location is the critical variable, could oscillations arise from competition between neurons representing any pair of locations or are they specific to fovea and periphery?

#### 4.4.4 A Potential Mechanism

The oscillatory activity observed in this study could arise from many sources including intracellular processes, interactions among neurons in IT and area-to-area interactions. None of these possibilities is clearly to be preferred to the others. However, it is worth noting that simple interactions at the network level, either confined to IT or involving other areas, could produce the effects described here in a comparatively straightforward manner. Networks of fatiguing neurons with reciprocal inhibitory connections are well known to give rise to oscillatory activity (Wilson et al., 2000). Furthermore, the ability of such networks to produce oscillatory activity qualitatively like that observed in our study can be demonstrated in terms of a model incorporating only two neurons, one responsive to the object and the other to the flanker, which give exponentially fatiguing responses and which inhibit each other (Figure 45A). These can be thought of as pyramidal neurons inhibiting each other via inhibitory interneurons not explicitly represented in the figure. We modeled these neurons as simple nodes with linear activation functions subject to fatigue. The state of neuron  $i$  ( $i = 1, 2$ ) was characterized in terms of three time-dependent variables:  $N_i(t)$  (net input),  $F_i(t)$  (level of fatigue) and  $O_i(t)$ , output, constrained by the following equations:

$$N_i(t) = w_e[E_i(t)] + w_i[I_i(t)];$$



$$d/dt F_i(t) = \tau^{-1} \{m[N_i(t)] - F_i(t)\};$$

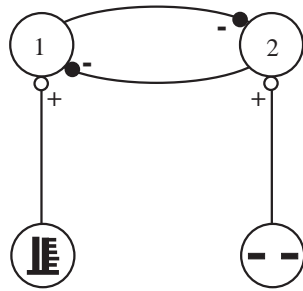
$$O_i(t) = N_i(t) - F_i(t) \text{ if } N_i(t) - F_i(t) > 0 \text{ else } O_i(t) = 0;$$

where  $w_e$  and  $w_i$  were the weights of excitatory and inhibitory synapses respectively,  $\tau$  was the time-constant of fatigue and  $m$  ( $0 \leq m \leq 1$ ) determined the strength of fatigue at asymptote.  $E_i(t)$ , the excitatory input to neuron  $i$ , was set to one or zero as the visual stimulus for which that neuron was selective was turned on or off.  $I_i(t)$ , the inhibitory input to neuron  $i$ , was set to the value of the other neuron's output. In the simulation depicted in Figure 45, the values of the constants were:  $w_e = 40$ ,  $w_i = 20$ ,  $\tau = 100$  ms and  $m = 0.7$ . The histograms of Figure 45B-D represent the activity of the object-selective neuron (neuron 1 in Figure 45A) under object-alone (Figure 45B), flanker-then-object (Figure 45C) and object-then-flanker (Figure 45D) conditions. Under the flanker-alone condition, this neuron was not active. We conclude that neurons adapting to visual stimulation at roughly the rate at which IT neurons adapt (Figure 45B) can give rise, through mutual inhibition, to oscillations at around 5 Hz and at the phases observed in this experiment.

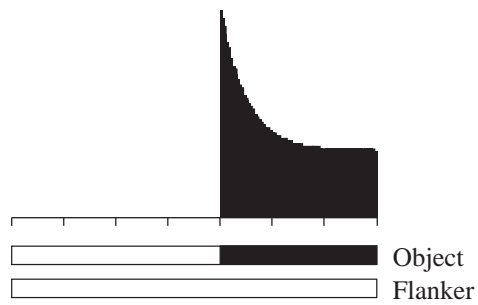
#### 4.4.5 Relation to Biased Competition

Desimone and his colleagues (Desimone and Duncan, 1995; Reynolds et al., 1999) have put forward a model of competitive effects in IT and other visual areas according to which simultaneously presented visual stimuli compete for neuronal representation. Competition in this model is proposed to arise from a combination of excitation and shunting inhibition such that two active afferents elicit a response intermediate in strength between the responses that those afferents would elicit

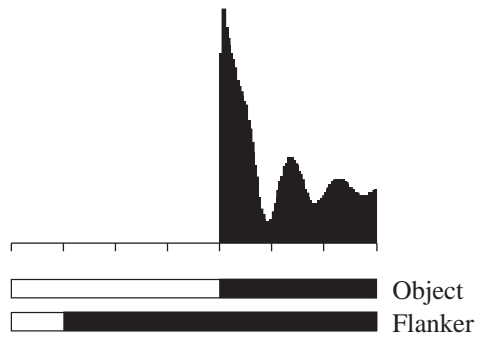
### A. Model



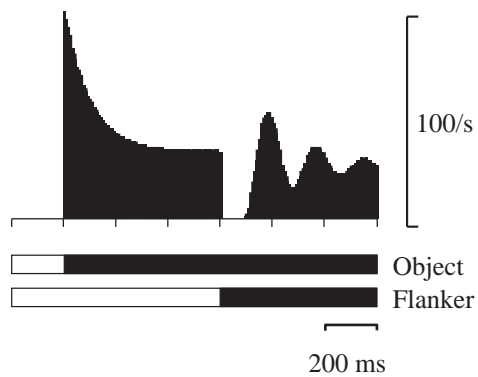
### B. Object Alone



### C. Flanker then Object



### D. Object then Flanker



**Figure 45. A simple model of oscillatory activity arising from reciprocal inhibition between neurons selectively responsive to different visual stimuli and subject to fatigue.** A. Neurons 1 and 2 selectively responsive to the object and the flanker respectively. B-D. Responses of neuron 1 under object-alone, flanker-then-object and object-then-flanker conditions.

independently (Reynolds et al., 1999). We will refer to this as an 'averaging' response. It would be parsimonious to assume that the same competitive mechanism underlies both averaging responses and oscillatory responses. However, the mechanism based on shunting inhibition (Reynolds et al., 1999) has no obvious potential to produce oscillatory responses because it involves no feedback loop. This observation raises the question of whether a network-based (as distinct from dendrite-based) mechanism could account for averaging responses and, if so, whether oscillatory responses could arise as an emergent property of the network. The idea that biased competition might involve reciprocal inhibition has been put forward before (Usher and Niebuhr, 1996; Deco and Lee, 2002). Moreover, inhibitory circuits are capable of normalizing each neuron's visual response to the response of the population as a whole (Carandini and Heeger, 1994; Carandini et al., 1997). In a normalizing network, two neurons responsive to stimulus A and stimulus B respectively respond at an intermediate level to the combination of A and B. Thus each neuron gives an averaging response to the combination of A and B. If averaging responses indeed depend on reciprocal inhibition, then the underlying inhibitory circuits might incidentally give rise to oscillatory activity as in the simple model of Figure 13.

#### **4.5 Summary**

In conclusion, we have demonstrated that low frequency oscillations in IT are enhanced when a foveal image is presented against the backdrop of an already present peripheral flanking display. We also found that turning on the peripheral display against the backdrop of an already present foveal stimulus elicits an initially negative oscillatory response. Finally, we observed that the occurrence of these phenomena is not critically

dependent on task context in that they occur both in a context promoting attention to the visual field periphery and in one requiring only central fixation. We suggest that these oscillations arise through mutually inhibiting populations of neurons responding to competing stimuli in the visual scene.

## **Chapter 5**

### **General Discussion**

#### **5.1 Summary**

The goal of the research presented in this thesis was to use single cell recording in the awake monkey to shed light on three specific outstanding issues concerning the role of IT in object processing. In Chapter 2, we demonstrated that neurons in IT differentiate less effectively between lateral mirror images than vertical mirror images, a phenomenon that parallels and may constitute the neural correlate of the behavioral confusion exhibited in animals and humans. In Chapter 3, we demonstrated that orientation discrimination training leads to significant increases in the selectivity of neurons for trained compared to untrained images, supporting the idea that experience-based changes in perception are paralleled by changes in the response properties of IT neurons. In Chapter 4, we characterized low-frequency oscillations in the activity of IT neurons and showed that they are enhanced by the addition of a second stimulus into the visual scene. We suggest that this phenomenon may result from mutual inhibition of competing populations of neurons.

The purpose of this chapter is to summarize the significance of these results. The following sections briefly review the findings of the three experiments and consider their relevance to our understanding of the contributions of IT to object perception. Outstanding questions raised by the results and directions for future study will be

discussed. More detailed discussions of the specific results are provided in the corresponding chapters.

## **5.2 Neurons in IT Exhibit Lateral Mirror Image Confusion**

In Chapter 2, we demonstrated that neurons in IT exhibit lateral mirror image confusion. Specifically, we presented mirror image stimuli to the fixating monkey while recording the activity of IT neurons and showed that neurons differentiated less effectively between lateral mirror images than vertical mirror images. This was true for images presented both foveally and peripherally. In addition, we observed that signals differentiating lateral and vertical mirror images arise early in the visual response, suggesting that this effect is due to intra-areal or feed-forward connections. This phenomenon parallels and may constitute a mechanism for lateral mirror image confusion as observed in behavior.

One question raised by these findings is how does lateral mirror image confusion in IT arise? As discussed in Chapter 2, one theory suggests that visual cortex in the two hemispheres are connected such that a neuron in one hemisphere preferring a given shape receives input from a neuron in the opposite hemisphere preferring the lateral mirror image of that shape (Corballis and Beale, 1970). Since information in the ipsilateral portion of IT receptive fields is received from neurons in the opposite hemisphere via the corpus callosum (Gross et al., 1977), we were able to directly test this theory by asking whether preference for members of lateral mirror image pairs was reversed in the two hemifields. While we observed a trend toward a reversed preference, the existence of lateral mirror image confusion for peripherally as well as foveally presented stimuli

meant that there were very few neurons showing a preference in either visual hemifield. Thus, the question of whether inter-hemispheric connectivity underlies lateral mirror image confusion in IT remains unresolved. Another way in which lateral mirror image confusion in IT neurons might arise would be via input from earlier cortical areas. For example, neurons within V4, or within IT itself, preferring opposite members of a lateral mirror image pair might converge onto the same neuron in IT, leading that IT neuron to respond equally to either image. However, it might also be the case that lateral mirror image confusion arises at even earlier stages in the ventral visual stream. Pasupathy and Connor (2001) show an example of a V4 neuron that responds more similarly to lateral mirror images than to vertical mirror images of a stimulus with three convex contours (their Figure 2, top center section). It would be of interest in future studies to determine whether this is a consistent effect in V4 neurons.

The results of the experiments described in Chapter 2 suggest that the selectivity of IT neurons does not simply reflect the physical similarity (or dissimilarity) between stimuli, since lateral and vertical mirror images are equally similar in terms of any isotropic measures. Rather, these findings suggest that the activity of IT neurons is related to the similarity of stimuli as perceived by the animal. This idea is supported by two previous studies that show that IT neurons tend to respond comparably to images that are found to be perceptually similar. In the first, Op de Beeck et al. (2001) demonstrated that the ability of monkeys to discriminate parameterized shapes and the selectivity of IT neurons for those same shapes were highly congruent, and that they deviated consistently from the parametric configurations. Miyashita et al. (1993) carried out a comparable study in which they examined the selectivity of IT neurons for fractal images that human

observers rated for similarity. The results showed that there was a tendency for neurons to respond similarly for stimuli that were rated as highly similar. Although our findings are consistent with the results from these two studies, the degree to which the animal confused mirror images was not explicitly measured in the study described in Chapter 2. While there is extensive evidence that monkeys do exhibit lateral mirror image confusion (Hamilton and Tieman, 1973; Riopelle et al., 1964), to confirm that responses of IT neurons for mirror images is related to the perception of those images would require a direct comparison between the degree to which the animals confuse mirror images and the degree to which neurons confuse mirror images. An analysis in Chapter 3 revealed a significant correlation between neuronal selectivity for and behavioral confusion of mirror images in monkeys performing a delayed match to sample task on shape orientations. This analysis is, however, inadequate for addressing the present question for two reasons. One, 180° rotated pairs were included in this analysis and therefore it does not offer a direct comparison between lateral and vertical mirror images. Two, the monkeys in that experiment were highly over-trained to discriminate mirror image stimuli. Therefore, they might have developed unique strategies to accomplish the discrimination and thus the behavior may not simply reflect the perceptual confusion of mirror images. To obtain an unbiased measure of behavioral confusion would require evaluating perceptual confusion in monkeys that have never been trained on an orientation discrimination task.

The question remains as to whether this phenomenon has any functional significance. Why would it be necessary for neurons in IT to equate lateral mirror images but not verticals? One theory posed by Gross and Bornstein (1978) offers a possible



answer to this question. Their theory suggests that lateral mirror image confusion arises as an adaptive phenomenon rather than a failure of the visual system. They suggest that since lateral reversals observed in nature almost always result from changes in viewpoint, they offer little information for the purposes of recognition. Therefore, since IT is thought to carry out processes underlying the recognition and perception of objects, neurons in this region would not need to encode such reflections. In contrast, vertical reversals rarely come about from changes in viewpoint and therefore do convey important information about the identity of the object. Thus, it might be useful for IT neurons to encode such reflections.

It is unclear, however, whether this adaptation would arise via evolutionary or developmental (experience-based) processes. If lateral mirror image confusion arises from developmental processes, then it might be expected that the confusion observed in IT neurons would be subject to modification by visual experience. If so, learning to discriminate between lateral mirror images may lead to an increase in selectivity for those images. We tested this in an experiment described in Chapter 3 and found that training to discriminate between both lateral and vertical mirror images leads to an overall increase in selectivity for both types of pairs, but no change in the relative selectivity for lateral vs. vertical pairs. It remains a question whether selective training on lateral mirror image discriminations alone would lead to a change in the relative selectivity of IT neurons for lateral and vertical mirror images.

Finally, if lateral mirror image confusion arises developmentally through an adaptive attenuation of activity reflecting accidental changes in viewpoint, then it might be the case that the degree to which IT neurons exhibit lateral mirror image confusion

varies depending on the stimulus. Specifically, it could be that objects typically seen in a canonical upright position (e.g., an automobile) may give rise to greater lateral mirror image confusion in IT neurons than an object that does not have a canonical upright (e.g., a house key). It would be of interest to determine whether lateral mirror image confusion as measured by neuronal selectivity and behavior, varies according to the orientation in which an object typically is viewed.

### **5.3 Training Increases the Selectivity of IT Neurons**

In Chapter 3, we demonstrated that orientation discrimination training significantly increased the selectivity of IT neurons for trained orientations. Specifically, we trained monkeys to discriminate among four orientations of each of ten shapes. We then recorded from IT neurons while monkeys (1) performed a DMS task with trained images and (2) passively viewed orientations of trained and untrained shapes. We found that training to discriminate shape orientation did not lead to changes in the response strength for preferred stimuli. However, training did result in significant increases in the selectivity of IT neurons for trained orientations, increases which took the form of small changes in many neurons. Furthermore, neuronal selectivity for orientations of trained images was significantly correlated with the monkeys' ability to discriminate these orientations.

These results support the idea that the experience-based increases in the ability to discriminate complex objects are related to changes in neuronal responses within IT for those objects. The data presented here are particularly important in light of the inconsistency of effects reported in previous studies. For example, Kobatake et al. (1998)

claimed that discrimination training leads to increases in response strength and selectivity for trained stimuli, but this conclusion is questionable as comparisons were made across different animals. Logothetis and Pauls (1995) claimed that discrimination training leads to dramatic increases in the selectivity of a few IT neurons for trained stimuli; however, their anecdotal observations were unsupported by any statistical quantitative analysis. In contrast, Baker et al. (2002) demonstrated that while discrimination training in a feature conjunction task did not result in changes in the overall strength of visual responses, it did lead to increases in the selectivity of neurons for trained stimuli. However, the findings of Baker et al. could conceivably be specific to the unique task conditions and stimuli used. The results of our study not only confirm those of Baker et al. but further extend them to the particular case of orientation discrimination. Ours is the first study demonstrating that changes in the selectivity of IT neurons result from extended training on an orientation discrimination task.

One issue raised in the present study concerns the form in which experience-based changes in selectivity take in IT neurons. In particular, the changes in selectivity as a result of orientation discrimination training, although significant, were small and distributed across large numbers of neurons. This is consistent with findings of Baker et al. (2002) who also demonstrated that significant increases in selectivity for trained images resulted from subtle changes across many neurons. Kobatake et al. (1998) also suggested that visual discriminations relied on responses of large populations of neurons with broadly tuned responses.

A question not addressed in the present study is on what timescale these training effects occur. Most studies that have examined the effects of discrimination training on

IT responses have done so in animals trained for several months (Logothetis and Pauls, 1995; Kobatake et al., 1998; Baker et al., 2002). It could be that the observed effects actually develop on a much shorter timescale. Indeed, Messinger et al. (2001) demonstrated that the effects of pair-association training were evident in the responses of IT neurons on the order of hours. Within a single training session, stimuli repeatedly paired began to evoke more similar responses from neurons. Moreover, the degree to which neurons exhibited pair-coding paralleled the performance of the animal.

It may be, however, that the neuronal mechanisms underlying association learning differ from those involved in discrimination learning. Erickson et al. (2000) examined the short-term effects of visual discrimination training on neuronal selectivity in perirhinal cortex. They found no change in the selectivity of neurons for trained stimuli compared to novel stimuli after one day of discrimination training. Therefore, it remains unclear how rapidly discrimination training effects might develop in IT. Although it would be of interest to investigate the time-course of neuronal changes during orientation discrimination training, there are some confounding issues. For one, animals may not sufficiently learn these discriminations in a single session to allow determining the time-course of selectivity changes in a single neuron. Any analysis of across-session changes would require comparisons between neurons. Further, as discussed above, the overall changes in selectivity in a single neuron were subtle and therefore may not lead to observable effects in a single session.

Finally, the experiments in Chapter 3 do not address the question of whether effects of training in IT could be accompanied by changes in other areas. The fact that we do see changes in response properties of IT neurons does not preclude the

involvement of other regions of cortex. In particular, there is strong evidence from lesion studies in humans and animals that the posterior parietal cortex is critical for performing discriminations of shape orientation (Eacott and Gaffan, 1991; Cooper and Humphreys, 2000; Harris et al., 2001). Furthermore, neurons in parietal area LIP have been shown to respond selectively to complex shapes (Serenio and Maunsell, 1998), and also to exhibit selectivity for the orientation of complex shapes (Rollenhagen and Olson, 2001). Future studies will address the question as to whether learning induced changes in neuronal selectivity occur in posterior parietal cortex.

#### **5.4 Low-Frequency Oscillations in IT**

In the studies described in Chapter 4, we quantified low-frequency oscillations observed in the activity of IT neurons and demonstrated that they are enhanced with the addition of a second stimulus. Specifically, we examined the activity of IT neurons in response to the presentation of multiple stimuli, a central pattern that excited the neuron and a peripheral stimulus that did not. We found that when the central pattern was presented in isolation, weak oscillatory activity was sometimes elicited at a frequency of 5-6 Hz. The tendency for visual responses to contain an oscillatory component increased dramatically when one stimulus was presented against the backdrop of another already present. Moreover, the phase of the oscillations was determined by which stimulus was presented first. In particular, we observed that the onset of the central pattern in the presence of the peripheral stimulus elicited a strong response with a marked oscillatory component phase-locked to pattern onset. Onset of the peripheral stimulus in the

presence of the central pattern elicited, in contrast, a succession of inhibitory troughs phase-locked to stimulus onset.

These results are significant in that they provide insight into the nature of the neural circuitry underlying visual responses of IT neurons. We suggest that this phenomenon may arise from mutual inhibition of competing populations of neurons in IT. In particular, it might be the case that neurons responsive to an image are initially excited by it, then succumb through fatigue to suppression by other neurons not responsive to the image, then recover and fire again, and so on. This idea is supported by the demonstration that networks consisting of fatiguing neurons with reciprocal inhibitory connections give rise to oscillatory activity (Wilson et al., 2000). If oscillatory activity in IT results from competitive interactions among stimuli as suggested above, then GABAergic interneurons demonstrated to mediate the presence of intrinsic inhibitory connections in IT (Wang et al., 2000) may be involved in this process.

Several questions are raised by the results of the studies presented in Chapter 4. First, if oscillations arise from mutual inhibition of competing populations of neurons, one might ask why they were observed, in weak form, when images were presented in isolation. One reason might be that within a single complex shape, there are separate features that activate separate, mutually inhibiting populations of neurons. Alternatively, it might be the case that the baseline activity of neurons not activated by the stimulus is sufficient to inhibit the activated populations of neurons to the degree that oscillations are invoked. Development of a more sophisticated computational model would help in further investigating the conditions that give rise to these oscillations.

A second question raised by the study described in Chapter 4 is whether oscillations would result from the onset of simultaneously presented stimuli. It might be that strong oscillatory activity occurs when one neuronal population is placed at a competitive disadvantage (through fatigue) by presenting its preferred stimulus first. In the present experiment, stimuli were presented either in isolation or in the presence of an already present stimulus. While oscillations have not been reported as a consequence of presenting multiple stimuli simultaneously (Miller et al., 1993a; Missal et al., 1999; Rolls and Tovee, 1995; Sato, 1989, 1995), it might be that the neurons tested in these studies did not have the tendency to oscillate. To determine whether staggered onset of stimuli is necessary for eliciting robust oscillatory activity would require assessing responses to the simultaneous onset of stimuli that elicit oscillatory responses when one is turned on before the other.

Another issue not resolved in the present study is whether oscillations arise from competition between neurons representing different images or representing different locations. Although IT receptive fields are generally thought to be quite large, recent evidence suggests that some receptive fields are as small as 3° of visual angle (Op de Beeck and Vogels, 2000). There also tends to be variability in response strength of neurons depending on where in the receptive field stimuli are presented (Op de Beeck and Vogels, 2000). These findings demonstrate that, to some degree, location is encoded in IT. Thus, it could be that neurons representing different locations in the visual field compete and give rise to the observed oscillations. It would be of interest to manipulate both the stimulus parameters and presentation locations to determine whether either or both are important in giving rise to the observed phenomenon.

## 5.5 Conclusions

In summary, the results of the studies described in this thesis support the role of IT in object perception and recognition. In particular, they suggest that the activity of neurons in IT reflects not just the physical properties of stimuli in the environment, but also how the animal perceives these stimuli. These results also support the idea that changes in the selectivity of IT neurons may provide the neural substrate for experience-based changes in perception. Finally, the characteristics of oscillatory activity in IT suggest that the representations of complex stimuli in IT are involved in dynamic competitive networks.



## APPENDIX

## **Appendix**

### **General Methods**

#### **A.1 Introduction**

This section explains the technical details of the procedures used to carry out experiments discussed in Chapters 2-4. All procedures described here were approved by the Carnegie Mellon University Animal Care and Use Committee and conformed to guidelines set forth in the United States Public Health Service Guide for the Care and Use of Laboratory Animals. Procedure details that are specific to a given experiment are described in the corresponding chapter.

#### **A.2 Chair Training**

At the onset of the experiment, animals were trained to leave their cages, climb into primate chairs and to be comfortable with their heads upright and restrained by a neck plate. The amount of time the animals spent in the primate chairs was gradually increased until animals were able to sit in the chairs for up to three hours without showing any signs of anxiety. Over this training period, animals were given free access to water while in the primate chairs. Access to water while the animals were in their cages was gradually restricted until they only received water while in the chairs. On days

the animals were not being trained, the same or more water was given in their cages that they earned during training.

### **A.3 Pedestal Implantation and Care**

Prior to training animals to perform the experimental tasks, surgery was carried out to (1) implant scleral search coils necessary to monitor eye movements and (2) affix head restraint bars to the monkeys' heads necessary to stabilize the head during electrophysiological recording. The surgery was carried out under aseptic conditions. The animals were given atropine (0.4 mg/kg, i.m.) followed by ketamine hydrochloride (20 mg/kg, i.m.) and valium (1.0 mg/kg, i.m.) to provide analgesia during preparation for surgery. Animals were maintained on gas anesthesia (isoflurane, 1-2%) throughout the surgery. To affix the head restraint bars, the scalp skin was first incised at the midline and the skull surface was exposed by retracting the muscles and removing the periosteum. Titanium bone-screws were then implanted around the rim of the exposed skull. Rapidly hardening acrylic was built up around the heads of the skull screws so as to completely cover the exposed skull. A plastic rod for attachment to the head-restraint clamp was then embedded in the acrylic pedestal. To implant scleral search coils, the conjunctival membranes were first resected near the limbus and a scleral search coil was implanted around the globe of each eye. The leads from each coil were run subcutaneously to a plug on the acrylic pedestal.

Immediately after the surgery, the animals were given Butorphanol (0.05 mg/kg, i.m.) to control postsurgical pain. Following surgery, if the animal showed any signs of pain or discomfort, such as lethargy or lack of appetite or thirst, additional injections of

butorphanol were administered as needed. During the week following the surgery, the monkey was given free access to food and water.

#### **A.4 Placement of Recording Chamber**

The site of chamber placement was chosen by considering both standard Horsley-Clarke coordinates reported in the literature (approximately A13 and L22), and brain images obtained through magnetic resonance imaging (MRI). Structural MR images were acquired for each monkey through the use of the Brüker 4.7 T magnet at the Pittsburgh NMR Center. Fiducial marks made visible in the images by means of a contrast agent aided in the placement of the recording chamber.

The surgery to place the recording chamber was carried out under aseptic conditions. The animals were given atropine (0.4 mg/kg, i.m.) followed by ketamine hydrochloride (20 mg/kg, i.m.) and valium (1.0 mg/kg, i.m.) throughout surgery. A 2 cm diameter disk of acrylic and skull was removed, leaving a cranial hole just large enough to accommodate the recording chamber. The chamber was placed at the appropriate location, and cemented into the hole with its base just above the exposed dural membrane. Routine measures designed to prevent infection were carried out at the beginning of each recording session or every 2-3 days when recording was not in progress.

## A.5 Electrophysiological Recording Methods

At the beginning of each recording session, a 23-gauge stainless steel sharpened guide tube was introduced into the cortex until its tip was approximately 1 cm below the surface of the cortex. Penetrations could be placed at 1 mm intervals by means of a nylon grid held rigidly in the recording chamber (Crist, 1988). A varnish coated tungsten electrode with an impedance of 0.5-8 M $\Omega$  at 1KHz (model #UEWLFCSEENIE, Frederick Haer Co., Bowdoinham, Maine) was advanced through the guide tube by means of a hydraulic microdrive (716-S, Narashige, Tokyo, Japan). Signals from the electrode were passed through an amplifier to a waveform analysis system (8701 Waveform Discriminator, Signal Processing Systems, Prospect, Australia) utilizing software which runs in real time on an independent Pentium-based platform. The system stores examples of action potentials generated by the neuron under study and then accepts or rejects each subsequent deflection of the trace as determined by a template matching algorithm. Action potentials determined to be generated by the neuron under study were transmitted as pulses to the data collection computer and stored with a temporal resolution of 1 ms.

Visual stimuli were displayed on a 14 inch video monitor placed at a distance of 38 cm from the monkey. Eye position was monitored through the use of a scleral search coil system (Robinson, 1963; Rempel, 1984) provided by Riverbend Instruments (Riverbend Instruments, Inc., Birmingham, AL). Reward in the form of approximately 0.1cc of juice was delivered through a spigot under control of a solenoid valve upon successful completion of each trial.

All aspects of the experiments including monitoring of neural activity, eye position, monkey's manual responses, generation and display of visual stimuli and delivery of reward were under on-line control by a Pentium-based computer. This data collection computer was equipped with add-on video and IO boards and ran "Cortex" software which was provided by Dr. Robert Desimone of the National Institute of Mental Health.

#### **A.6 Data Display, Storage and Analysis**

Data were stored permanently on compact disks. Off-line analysis was carried out on a pentium-based computer. During off-line analysis, the data were viewed in histogram and raster format, with rasters aligned on the occurrence of a recorded event. Statistical analysis was carried out using Statistica (Statsoft, Inc., Tulsa, OK) and custom-written routines in Matlab (Mathworks, Natick, MA). Methods for analyzing data from individual neurons are described above in connection with individual experiments.

#### **A.7 Localization of Recording Sites**

During data collection, the mediolateral and anterior-posterior coordinates of each vertical track were noted relative to the 1 cm square grid centered on the recording chamber. The depth of recording sites was noted relative to the base of the grid, to the bone shelf ventral to the temporal lobe, and to white and grey matter as identified through audio monitor output and signals on the oscilloscope. The location of recording sites was ascertained by analysis of structural MR images. A plexiglass cylinder with columns of

contrasting agent at selected locations within the chamber was inserted into the chamber just prior to scanning. This allowed for the translation of the grid and depth coordinates into brain coordinates by showing the brain relative to MR-visible fiducial markers placed at known grid locations. Frontoparallel sections of 2 mm thickness spanning the entire brain were collected. The locations of recording sites are given in Horsley-Clarke coordinates defined in millimeters relative to the interaural plane.

## BIBLIOGRAPHY



## Bibliography

- Aggleton, J. P., Burton, M. J., & Passingham, R. E. (1980). Cortical and subcortical afferents to the amygdala of the rhesus monkey (*Macaca mulatta*). *Brain Res*, 190(2), 347-68.
- Albright, T. D., & Gross, C. G. (1990). Do inferior temporal cortex neurons encode shape by acting as fourier descriptor filters. *Proceedings of the International Conference on Fuzzy Logic and Neural Networks*, 2, 375-378.
- Baker, C. I., Behrmann, M., & Olson, C. R. (2002). Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat Neurosci*, 5(11), 1210-6.
- Baylis, G. C., & Rolls, E. T. (1987). Responses of neurons in the inferior temporal cortex in short term and serial recognition memory tasks. *Exp Brain Res*, 65(3), 614-22.
- Booth, M. C., & Rolls, E. T. (1998). View-invariant representations of familiar objects by neurons in the inferior temporal visual cortex. *Cereb Cortex*, 8(6), 510-23.
- Bornstein, M. H., Gross, C. G., & Wolf, J. Z. (1978). Perceptual similarity of mirror images in infancy. *Cognition*, 6(2), 89-116.
- Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci*, 13(1), 87-100.
- Carandini, M., & Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science*, 264(5163), 1333-6.
- Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *J Neurosci*, 17(21), 8621-44.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (1993). A neural basis for visual search in inferior temporal cortex [see comments]. *Nature*, 363(6427), 345-7.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol*, 80(6), 2918-40.
- Cheng, K., Saleem, K. S., & Tanaka, K. (1997). Organization of corticostriatal and corticoamygdalar projections arising from the anterior inferotemporal area TE of

- the macaque monkey: a Phaseolus vulgaris leucoagglutinin study. *J Neurosci*, 17(20), 7902-25.
- Colby, C. L., & Duhamel, J. R. (1996). Spatial representations for action in parietal cortex. *Brain Res Cogn Brain Res*, 5(1-2), 105-15.
- Connors, B. W., & Gutnick, M. J. (1990). Intrinsic firing patterns of diverse neocortical neurons. *Trends Neurosci*, 13(3), 99-104.
- Cooper, A. C., & Humphreys, G. W. (2000). Task-specific effects of orientation information: neuropsychological evidence. *Neuropsychologia*, 38(12), 1607-15.
- Corballis, M. C., & Beale, I. L. (1970). Bilateral symmetry and behavior. *Psychol Rev*, 77(5), 451-64.
- Cowey, A., & Gross, C. G. (1970). Effects of foveal prestriate and inferotemporal lesions on visual discrimination by rhesus monkeys. *Exp Brain Res*, 11(2), 128-44.
- Crist, C. R., Yamasaki, D. S. G., Komatsu, H., & Wurtz, R. H. (1988). A grid system and a microsyringe for single cell recording. *J. Neurosci. Methods*, 26, 117-122.
- Cutzu, F., & Edelman, S. (1998). Representation of object similarity in human vision: psychophysics and a computational model. *Vision Res*, 38(15-16), 2229-57.
- Dean, P. (1976). Effects of inferotemporal lesions on the behavior of monkeys. *Psychol Bull*, 83(1), 41-71.
- Dean, P. (1982). Visual behavior in monkeys with inferotemporal lesions. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 587-628). Cambridge, MA: MIT Press.
- Deco, G., & Lee, T. S. (2002). A unified model of spatial and object attention based on inter-cortical biased competition. *Neurocomputing*, 44-46, 775-781.
- Desimone, R., Fleming, J., & Gross, C. G. (1980). Prestriate afferents to inferior temporal cortex: an HRP study. *Brain Res*, 184(1), 41-55.
- Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *J Neurosci*, 4(8), 2051-62.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu Rev Neurosci*, 18, 193-222.
- Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philos Trans R Soc Lond B Biol Sci*, 353(1373), 1245-55.
- Eacott, M. J., & Gaffan, D. (1991). The role of monkey inferior parietal cortex in visual discrimination of identity and orientation of shapes. *Behav Brain Res*, 46(1), 95-8.

- Erickson, C. A., & Desimone, R. (1999). Responses of macaque perirhinal neurons during and after visual stimulus association learning. *J Neurosci*, 19(23), 10404-16.
- Erickson, C. A., Jagadeesh, B., & Desimone, R. (2000). Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. *Nat Neurosci*, 3(11), 1143-8.
- Felleman, D. J., & Essen, D. C. V. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1(1), 1-47.
- Fujita, I., Tanaka, K., Ito, M., & Cheng, K. (1992). Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, 360(6402), 343-6.
- Fuster, J. M., & Jervey, J. P. (1981). Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science*, 212(4497), 952-5.
- Fuster, J. M. (1990). Inferotemporal units in selective visual attention and short-term memory. *J Neurophysiol*, 64(3), 681-97.
- Gibson, J. R., & Maunsell, J. H. (1997). Sensory modality specificity of neural activity related to memory in visual cortex. *J Neurophysiol*, 78(3), 1263-75.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25.
- Gross, C. G., Bender, D. B., & Rocha-Miranda, C. E. (1969). Visual receptive fields of neurons in inferotemporal cortex of the monkey. *Science*, 166(910), 1303-6.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, 35, 96-111.
- Gross, C. G., Bender, D. B., & Mishkin, M. (1977). Contributions of the corpus callosum and the anterior commissure to visual activation of inferior temporal neurons. *Brain Res*, 131(2), 227-39.
- Gross, C. G., & Bornstein, M. H. (1978). Left and right in science and art. *Leonardo*, 11, 29-38.
- Gross, C. G. (1978). Inferior temporal lesions do not impair discrimination of rotated patterns in monkeys. *J Comp Physiol Psychol*, 92(6), 1095-1109.
- Hamilton, C. R., & Tieman, S. B. (1973). Interocular transfer of mirror image discriminations by chiasm-sectioned monkeys. *Brain Res*, 64, 241-55.
- Harris, I. M., Harris, J. A., & Caine, D. (2001). Object orientation agnosia: a failure to find the axis? *J Cogn Neurosci*, 13(6), 800-12.

- Hasegawa, I., & Miyashita, Y. (2002). Categorizing the world: expert neurons look into key features. *Nat Neurosci*, 5(2), 90-1.
- Herzog, A. G., & Van Hoesen, G. W. (1976). Temporal neocortical afferent connections to the amygdala in the rhesus monkey. *Brain Res*, 115(1), 57-69.
- Huttenlocher, J. (1967). Children's ability to order and orient objects. *Child Dev*, 38(4), 1169-76.
- Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *J Neurophysiol*, 73(1), 218-26.
- Janssen, P., Vogels, R., & Orban, G. A. (2000). Selectivity for 3D shape that reveals distinct areas within macaque inferior temporal cortex. *Science*, 288(5473), 2054-6.
- Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, 71(3), 856-867.
- Kobatake, E., Wang, G., & Tanaka, K. (1998). Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J Neurophysiol*, 80(1), 324-30.
- Kovacs, G., Vogels, R., & Orban, G. A. (1995). Selectivity of macaque inferior temporal neurons for partially occluded shapes. *J Neurosci*, 15(3 Pt 1), 1984-97.
- Lehky, S. R., Tanaka, K. Macaque inferotemporal visual response during memory – intensive and passive viewing tasks. Program No. 160.10. 2002 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience, 2002. CD-ROM.
- Logothetis, N. K., & Pauls, J. (1995). Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb Cortex*, 5(3), 270-88.
- Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. *Curr Biol*, 5(5), 552-63.
- Messinger, A., Squire, L. R., Zola, S. M., & Albright, T. D. (2001). Neuronal representations of stimulus associations develop in the temporal lobe during learning. *Proc Natl Acad Sci U S A*, 98(21), 12239-44.
- Miller, E. K., Li, L., & Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. *Science*, 254(5036), 1377-9.

- Miller, E. K., Gochin, P. M., & Gross, C. G. (1993a). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Res*, 616(1-2), 25-9.
- Miller, E. K., Li, L., & Desimone, R. (1993b). Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *J Neurosci*, 13(4), 1460-78.
- Miller, E. K., & Desimone, R. (1994). Parallel neuronal mechanisms for short-term memory. *Science*, 263(5146), 520-2.
- Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., Bettucci, D., Mortara, F., Mutani, R., Terazzi, E., & Davidson, D. L. W. (1991). Perception and action in 'visual form agnosia'. *Brain*, 114, 405-428.
- Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.
- Missal, M., Vogels, R., Li, C. Y., & Orban, G. A. (1999). Shape interactions in macaque inferior temporal neurons. *J Neurophysiol*, 82(1), 131-42.
- Miyashita, Y., & Chang, H. S. (1988). Neuronal correlate of pictorial short-term memory in the primate temporal cortex. *Nature*, 331(6151), 68-70.
- Miyashita, Y., Date, A., & Okuno, H. (1993). Configurational encoding of complex visual forms by single neurons of monkey temporal cortex. *Neuropsychologia*, 31(10), 1119-31.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229(4715), 782-4.
- Murata, A., Gallese, V., Luppino, G., Kaseda, M., & Sakata, H. (2000). Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J Neurophysiol*, 83(5), 2580-601.
- Nakamura, K., Mikami, A., & Kubota, K. (1991). Unique oscillatory activity related to visual processing in the temporal pole of monkeys. *Neurosci Res*, 12(1), 293-9.
- Nakamura, K., Mikami, A., & Kubota, K. (1992). Oscillatory neuronal activity related to visual short-term memory in monkey temporal pole. *Neuroreport*, 3(1), 117-20.
- Nakamura, H., Gattass, R., Desimone, R., & Ungerleider, L. G. (1993). The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *J Neurosci*, 13(9), 3681-91.
- Naya, Y., Yoshida, M., & Miyashita, Y. (2003). Forward processing of long-term associative memory in monkey inferotemporal cortex. *J Neurosci*, 23(7), 2861-71.

- Op De Beeck, H., & Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *J Comp Neurol*, 426(4), 505-18.
- Op de Beeck, H., Wagemans, J., & Vogels, R. (2001). Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat Neurosci*, 4(12), 1244-52.
- Optican, L. M., & Richmond, B. J. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. III. Information theoretic analysis. *J Neurophysiol*, 57(1), 162-78.
- Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: position-specific tuning for boundary conformation. *J Neurophysiol*, 86(5), 2505-19.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47, 329-342.
- Rommel, R. S. (1984). An inexpensive eye movement monitor using the scleral search coil technique. *IEEE Trans. Biomed. Eng.*, 31, 388-390.
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci*, 19(5), 1736-53.
- Richmond, B. J., & Optican, L. M. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. II. Quantification of response waveform. *J Neurophysiol*, 57(1), 147-61.
- Richmond, B. J., Optican, L. M., Podell, M., & Spitzer, H. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics. *J Neurophysiol*, 57(1), 132-46.
- Riopelle, A. J., Rahm, U., Itoigawa, N., & Draper, W. A. (1964). Discrimination of mirror-image patterns by rhesus monkeys. *Percept Mot Skills*, 19, 383-389.
- Rollenhagen, J. E., & Olson, C. R. (2000). Mirror-image confusion in single neurons of the macaque inferotemporal cortex. *Science*, 287(5457), 1506-8.
- Rollenhagen, J.R. and Olson, C.R. Encoding of object orientation in monkey posterior parietal cortex. Program No. 399.5. 2001 Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience, 2001. CD-ROM.
- Rolls, E. T., Baylis, G. C., Hasselmo, M. E., & Nalwa, V. (1989). The effect of learning on the face selective responses of neurons in the cortex in the superior temporal sulcus of the monkey. *Exp Brain Res*, 76(1), 153-64.
- Rolls, E. T., & Tovee, M. J. (1995). The responses of single neurons in the temporal visual cortical areas of the macaque when more than one stimulus is present in the receptive field. *Exp Brain Res*, 103(3), 409-20.

- Rudel, R. G., & Teuber, H.-L. (1963). Discrimination of direction of line in children. *J comp physiol Psychol*, 56(5), 892-898.
- Sakai, K., & Miyashita, Y. (1991). Neural organization for the long-term memory of paired associates. *Nature*, 354(6349), 152-5.
- Sakai, K., & Miyashita, Y. (1994). Neuronal tuning to learned complex forms in vision. *Neuroreport*, 5(7), 829-32.
- Saleem, K. S., & Tanaka, K. (1996). Divergent projections from the anterior inferotemporal area TE to the perirhinal and entorhinal cortices in the macaque monkey. *J Neurosci*, 16(15), 4757-75.
- Saleem, K. S., Suzuki, W., Tanaka, K., & Hashikawa, T. (2000). Connections between anterior inferotemporal cortex and superior temporal sulcus regions in the macaque monkey. *J Neurosci*, 20(13), 5083-101.
- Sary, G., Vogels, R., & Orban, G. A. (1993). Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science*, 260(5110), 995-7.
- Sato, T., Kawamura, T., & Iwai, E. (1980). Responsiveness of inferotemporal single units to visual pattern stimuli in monkeys performing discrimination. *Exp Brain Res*, 38(3), 313-9.
- Sato, T. (1989). Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. *Exp Brain Res*, 77(1), 23-30.
- Sato, T. (1995). Interactions between two different visual stimuli in the receptive fields of inferior temporal neurons in macaques during matching behaviors. *Exp Brain Res*, 105(2), 209-19.
- Schwartz, E. L., Desimone, R., Albright, T. D., & Gross, C. G. (1983). Shape recognition and inferior temporal neurons. *Proc Natl Acad Sci U S A*, 80(18), 5776-8.
- Sekuler, R. W., & Houlihan, K. (1968). Discrimination of mirror-images: choice time analysis of human adult performance. *Q J Exp Psychol*, 20(2), 204-7.
- Seltzer, B., & Pandya, D. N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res*, 149(1), 1-24.
- Sereno, A. B., & Maunsell, J. H. (1998). Shape selectivity in primate lateral intraparietal cortex. *Nature*, 395(6701), 500-3.
- Serpell, R. (1971). Discrimination of orientation by Zambian children. *J Comp Physiol Psychol*, 75(2), 312-6.

- Sheinberg, D. L., & Logothetis, N. K. (1997). The role of temporal cortical areas in perceptual organization. *Proc Natl Acad Sci U S A*, 94(7), 3408-13.
- Sheinberg, D. L., & Logothetis, N. K. (2001). Perceptual learning and the development of complex visual representations in temporal cortical neurons. In M. Fahle & T. Poggio (Eds.), *Perceptual Learning* (pp. 95-124). Cambridge: MIT Press.
- Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, 415(6869), 318-20.
- Snyder, L. H., Batista, A. P., & Andersen, R. A. (2000). Intention-related activity in the posterior parietal cortex: a review. *Vision Res*, 40(10-12), 1433-41.
- Sugihara, T., Edelman, S., & Tanaka, K. (1998). Representation of objective similarity among three-dimensional shapes in the monkey. *Biol Cybern*, 78(1), 1-7.
- Sutherland, N. S. (1960). Visual discrimination of orientation by octopus: mirror images. *Brit J Psychol*, 51(1), 9-18.
- Suzuki, W. A. (1996). Neuroanatomy of the monkey entorhinal, perirhinal and parahippocampal cortices: Organization of cortical inputs and interconnections with amygdala and striatum. *Seminars in the neurosciences*, 8, 3-12.
- Tamura, H., & Tanaka, K. (2001). Visual response properties of cells in the ventral and dorsal parts of the macaque inferotemporal cortex. *Cereb Cortex*, 11(5), 384-99.
- Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J Neurophysiol*, 66(1), 170-89.
- Todrin, D. C., & Blough, D. S. (1983). The discrimination of mirror-image forms by pigeons. *Percept Psychophys*, 34(4), 397-402.
- Tomonaga, M., & Matsuzawa, T. (1992). Perception of complex geometric figures in chimpanzees (*Pan troglodytes*) and humans (*Homo sapiens*): analyses of visual similarity on the basis of choice reaction time. *J Comp Psychol*, 106(1), 43-52.
- Tovee, M. J., & Rolls, E. T. (1992). Oscillatory activity is not evident in the primate temporal visual cortex with static stimuli. *Neuroreport*, 3(4), 369-72.
- Tovee, M. J., Rolls, E. T., Treves, A., & Bellis, R. P. (1993). Information encoding and the responses of single neurons in the primate temporal visual cortex. *J Neurophysiol*, 70(2), 640-54.
- Tsunoda, K., Yamane, Y., Nishizaki, M., & Tanifuji, M. (2001). Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nat Neurosci*, 4(8), 832-8.



- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.
- Ungerleider, L. G., & Haxby, J. V. (1994). 'What' and 'where' in the human brain. *Current Opinion in Neurobiology*, 4, 157-165.
- Usher, M., & Niebuhr, E. (1996). Modeling the temporal dynamics of IT neurons in visual search: A mechanism for top-down selective attention. *J Cog Neurosci*, 8, 311-327.
- Van Hoesen, G., & Pandya, D. N. (1975). Some connections of the entorhinal (area 28) and perirhinal (area 35) cortices of the rhesus monkey. I. Temporal lobe afferents. *Brain Res*, 95(1), 1-24.
- Vogels, R., & Orban, G. A. (1994a). Activity of inferior temporal neurons during orientation discrimination with successively presented gratings. *J Neurophysiol*, 71(4), 1428-51.
- Vogels, R., & Orban, G. A. (1994b). Does practice in orientation discrimination lead to changes in the response properties of macaque inferior temporal neurons? *Eur J Neurosci*, 6(11), 1680-90.
- Vogels, R., Sary, G., & Orban, G. A. (1995). How task-related are the responses of inferior temporal neurons? *Vis Neurosci*, 12(2), 207-14.
- von Bonin, G., & Bailey, P. (1947). *The neocortex of Macaca mulatta*. Urbana: University of Illinois Press.
- Wallis, G., & Bulthoff, H. (1999). Learning to recognize objects. *Trends Cogn Sci*, 3(1), 22-31.
- Wang, Y., Fujita, I., & Murayama, Y. (2000). Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nat Neurosci*, 3(8), 807-13.
- Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1993). Subcortical connections of inferior temporal areas TE and TEO in macaque monkeys. *J Comp Neurol*, 335(1), 73-91.
- Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1994). Connections of inferior temporal areas TEO and TE with parietal and frontal cortex in macaque monkeys. *Cereb Cortex*, 4(5), 470-83.
- Wilson, H. R., Krupa, B., & Wilkinson, F. (2000). Dynamics of perceptual oscillations in form vision. *Nat Neurosci*, 3(2), 170-6.

- Xiang, J. Z., & Brown, M. W. (1998). Differential neuronal encoding of novelty, familiarity and recency in regions of the anterior temporal lobe. *Neuropharmacology*, 37(4-5), 657-76.
- Xiang, J. Z., & Brown, M. W. (1999). Differential neuronal responsiveness in primate perirhinal cortex and hippocampal formation during performance of a conditional visual discrimination task. *Eur J Neurosci*, 11(10), 3715-24.
- Yamane, S., Kaji, S., & Kawano, K. (1988). What facial features activate face neurons in the inferotemporal cortex of the monkey? *Exp Brain Res*, 73(1), 209-14.
- Zar, J. H. (1999). *Biostatistical Analysis* (4th ed.). Upper Saddle River: Prentice Hall.